

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 02:03:16 ; Search time 8287 Seconds
(without alignments)
12334.771 Million cell updates/sec

Title: US-10-092-390-1
Perfect score: 3423
Sequence: 1 atggttatttctttgaactc.....gcagcagcagcagtgaatga 3423

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*

28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	1334.2	39.0	3162	11	BC029999	BC029999 Homo sapi
2	992.6	29.0	2910	29	AY406554	AY406554 Homo sapi
3	972.6	28.4	3568	11	AK051642	AK051642 Mus muscu
4	933.2	27.3	2910	29	AY406556	AY406556 Mus muscu
5	892.6	26.1	3466	11	AK032661	AK032661 Mus muscu
6	653.6	19.1	2910	29	AY406555	AY406555 Pan trogl
7	646.2	18.9	3556	11	AK053551	AK053551 Mus muscu
8	587.6	17.2	789	14	CD803668	CD803668 UI-M-GV0-
9	539	15.7	921	13	BU215784	BU215784 603106167
10	530.4	15.5	755	14	CD802967	CD802967 UI-M-GV0-
11	508.6	14.9	643	13	BU056532	BU056532 UI-M-FO0-
12	470	13.7	2944	11	AK048840	AK048840 Mus muscu
13	456.6	13.3	755	12	BG828819	BG828819 602751395
14	433	12.6	565	12	BM719978	BM719978 UI-E-EJ0-
c 15	432	12.6	598	12	BM676825	BM676825 UI-E-EJ0-
16	417.4	12.2	846	13	BX741190	BX741190 BX741190
17	405.4	11.8	649	9	AL860947	AL860947 AL860947
18	398.8	11.7	798	13	BU112175	BU112175 603126134
19	379	11.1	538	10	BB762574	BB762574 BB762574
20	370	10.8	715	9	AI958909	AI958909 fd05g03.y
21	367.6	10.7	438	12	BM484594	BM484594 538509 MA
22	366.8	10.7	779	10	BF529240	BF529240 602041695
23	359.6	10.5	641	10	BB623063	BB623063 BB623063
24	359.6	10.5	744	13	BY741291	BY741291 BY741291
25	354.6	10.4	488	14	H15471	H15471 ym29b02.r1
26	352.8	10.3	734	14	CF745181	CF745181 UI-M-GV0-
27	328.6	9.6	772	14	CF538219	CF538219 UI-M-GH0-
28	328.2	9.6	584	10	BF686873	BF686873 602102830
29	318.8	9.3	694	13	BY734908	BY734908 BY734908
30	313.4	9.2	760	13	BY705363	BY705363 BY705363
31	309.4	9.0	1123	12	BM563529	BM563529 AGENCOURT
32	302	8.8	559	10	AW658138	AW658138 93921 MAR
33	301.2	8.8	563	14	CF538582	CF538582 UI-M-GI0-
34	296.2	8.7	937	10	BF180097	BF180097 601806458
35	294.4	8.6	852	9	AL040177	AL040177 DKFZp434F
36	293.2	8.6	401	10	BE664785	BE664785 152639 MA
37	283.4	8.3	451	14	CB786815	CB786815 AMGNNUC:N
38	282.2	8.2	463	10	AW445546	AW445546 81787 MAR
39	271.2	7.9	477	10	BF442243	BF442243 258883 MA
40	268.8	7.9	693	14	CF739003	CF739003 UI-M-HD0-
41	260	7.6	501	10	AW478869	AW478869 22097 MAR
42	255	7.4	634	13	BY726500	BY726500 BY726500
43	253.6	7.4	634	10	BB650216	BB650216 BB650216
44	251	7.3	843	13	BU282192	BU282192 603862751
45	250.2	7.3	600	14	CA893118	CA893118 B0177A01-

ALIGNMENTS

RESULT 1

BC029999

LOCUS BC029999 3162 bp mRNA linear HTC 06-MAY-2002

DEFINITION Homo sapiens, clone IMAGE:4156083, mRNA.

ACCESSION BC029999

VERSION BC029999.1 GI:20455873

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3162)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (06-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USAREMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: David N. Louis, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>Contact: amg@bcm.tmc.eduGunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 52 Row: a Column: 10

This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14192940

This clone has the following problem: no 5' EST match.

FEATURES

source

Location/Qualifiers

1. .3162

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4156083"

/tissue_type="Brain, anaplastic oligodendroglioma with
1p/19q loss"

/clone_lib="NCI_CGAP_Brn67"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 39.0%; Score 1334.2; DB 11; Length 3162;
Best Local Similarity 69.4%; Pred. No. 0;
Matches 1845; Conservative 0; Mismatches 808; Indels 6; Gaps 2;

```
Qy      74 CTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACTCAGTGACTGTGC 133
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     229 CCCTGAACCCCGAGGACCCCAACGTGTGCAGCCACTGGGAGAGCTATGCTGTGACTGTCC 288

Qy     134 AAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGCACTGACATTCTAA 193
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     289 AGGAATCGTATGCACACCCCTTCGATCAGATCTATTACACACGATGCACAGACATCCTCA 348

Qy     194 ACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTATCGACATGGGGAGA 253
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     349 ACTGGTTCAAGTGCACCAGGCACCGGATCAGTTATAAGACGGCGTATCGGAGAGGCCTCC 408

Qy     254 AGACTATGTATAGGCGCAAGTCTCAGTGTGTCTGGATTTTATGAAAGCGGGGAAATGT 313
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     409 GGACCATGTACCGGCGGAGGTCCCAGTGTGCTGCCCTGGCTACTATGAGAGCGGAGACTTCT 468

Qy     314 GTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCTCCAAACACCTGTC 373
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     469 GCATACCCCTGTGTACGGAGGAGTGTGTGCACGGCCGCTGCGTTTCCCCGGACACCTGCC 528

Qy     374 AGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGATGGTGATCACTGGG 433
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     529 ACTGCGAGCCTGGCTGGGGAGGGCCCCGACTGCTCCAGCGGCTGCGACAGCGACCACTGGG 588

Qy     434 GTCCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGCAACCCCATCACCG 493
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     589 GGCCCCCACTGCAGCAACCGGTGCCAGTGCCAGAACGGCGCCCTGTGTAACCCCATCACAG 648

Qy     494 GGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGCAGG 553
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     649 GCGCCTGCGTGTGCGCCGCCGGCTTCCGTGGATGGCGCTGCGAGGAGCTCTGCGCGCCTG 708

Qy     554 GCACCTATGGTAACGACTGTGCATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCGACC 613
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     709 GCACCCACGGCAAGGGATGCCAGCTGCCGTGCCAGTGCCGACACGGTGCCAGCTGCGACC 768

Qy     614 ACGTCACGGGGGAATGCCGCTGCCACCAGGATACACGGAGCCTTCTGTGAGGATCTTT 673
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     769 CCCGCGCCGGCGAGTGCCTCTGCGCACCTGGCTACACGGCGTCTACTGCGAGGAGCTGT 828

Qy     674 GTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAAATGGAGGAG 733
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     829 GCCCTCCTGGGAGCCATGGAGCTCACTGTGAGCTGCGCTGCCCCGTGTCAGAATGGGGGCA 888

Qy     734 TGTGTGCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGTGTG 793
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     889 CCTGCCACCACATCACTGGCGAGTGTGCCTGCCCCCAGGCTGGACGGGAGCAGTGTGTG 948

Qy     794 GTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCCATA 853
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     949 CCCAGCCCTGCCACCAGGGACATTTGGCCAGAAGTGCAGCCAGGATTGTCCTTGCCACC 1008

Qy     854 ATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAGGGG 913
```


Db	1009	ATGGAGGGCAGTGTGACCACGTGACTGGACAGTGCCACTGTACAGCTGGATACATGGGGG	1068
Qy	914	AACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGCC	973
Db	1069	ACAGGTGCCAAGAGGAGTGCCCTTCGGGTCTTCGGCTTCCAGTGCTCACAGCGCTGTG	1128
Qy	974	AGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAGGCT	1033
Db	1129	ACTGCCACAATGGGGGGCAGTGTTCACCCACCACGGGTGCCTGCGAGTGTGAGCCTGGCT	1188
Qy	1034	TTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTACGGCATCAAATGTG	1093
Db	1189	ACAAGGGCCCACGCTGCCAGGAGCGACTGTGCCCGGAGGGCCTGCATGGCCCAGGCTGCA	1248
Qy	1094	ACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTACCCCCATGTCTGGAGAGT	1153
Db	1249	CCCTGCCCTGCCCTGTGACGCTGACAACACCATCAGCTGCCACCCAGTAACTGGAGCTT	1308
Qy	1154	GTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGATTCT	1213
Db	1309	GTACCTGCCAGCCAGGCTGGTCTGGTCACCACTGCAATGAATCCTGCCCTGTTGGCTACT	1368
Qy	1214	ACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGA	1273
Db	1369	ATGGCGATGGCTGCCAGCTGCCTTGACCTGTCAGAATGGCGCCGACTGCCACAGCATCA	1428
Qy	1274	CTGGAAAGTGACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTC	1333
Db	1429	CTGGGGGCTGCACTTGTGCTCCGGGCTTCATGGGAGAGGTCTGTGCCGTTTCTGTGCAG	1488
Qy	1334	TGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCT	1393
Db	1489	CAGGGACCTATGGCCCCAACTGCTCGTCCATCTGTAGCTGTAACAATGGTGGCACCTGCT	1548
Qy	1394	CTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCA	1453
Db	1549	CCCCAGTAGATGGCTCCTGTACCTGCAAGGAAGGTGGCAGGGCCTGGACTGCACCCTGC	1608
Qy	1454	GATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTTAACATGCCAGTGCCTCAACGGGG	1513
Db	1609	CATGTCCCAGTGGGACGTGGGGCCTGAACTGCAACGAGAGCTGCACCTGTGCCAATGGGG	1668
Qy	1514	GAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAAT	1573
Db	1669	CAGCCTGCAGCCCCATAGACGGCTCCTGCTCCTGCACTCCTGGCTGGCTGGGAGACACCT	1728
Qy	1574	GCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCA	1633
Db	1729	GTGAGCTGCCTTGCCCGGATGGCACATTTGGGCTGAACTGCAGTGAACACTGTGACTGCA	1788
Qy	1634	GCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAG	1693
Db	1789	GCCATGCTGATGGATGTGACCCCGTCACAGGCCACTGCTGCTGCCTGGCCGGATGGACAG	1848
Qy	1694	GTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCT	1753

Db 1849 GCATCCGCTGTGACAGCACGTGTCCACCTGGCCGCTGGGGCCCCAACTGCTCTGTCTCCT 1908

Qy 1754 GCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAG 1813
 || ||| | |||| | || ||||| || |||| | ||||| |||

Db 1909 GCAGCTGTGAGAATGGAGGCTCCTGCTCCCCAGAGGATGGGAGCTGCGAGTGTGCCCTG 1968

Qy 1814 GCTTCCGAGGCACCACTTGTTCAGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCTGCA 1873
 ||||| || || |||| |||| ||||| || |||| || ||||

Db 1969 GCTTCCGAGGACCCTTATGCCAGAGAATCTGCCCCCTGGGTTCTATGGCCACGGCTGCG 2028

Qy 1874 GCCAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGT 1933
 |||| ||||| | |||| ||||| ||||| ||||| |||| |

Db 2029 CCCAGCCATGCCCCCTCTGCGTGCACAGCAGCAGGCCCTGCCACCACATCAGCGGCATCT 2088

Qy 1934 GTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGAT 1993
 |||| ||| || || || || || || || || || || || || || || ||

Db 2089 GTGAGTGCCTCCCAGGATTCTCTGGAGCTCTCTGCAACCAAGTGTGTGCTGGAGGATACT 2148

Qy 1994 TTGGGAAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTG 2053
 |||| | ||||| || || |||| ||||| |||| | || || |

Db 2149 TTGGGCAGGACTGTGCCAGCTCTGCTCCTGTGCCAACAACGGGACCTGCAGCCCTATCG 2208

Qy 2054 ACAGATCTTGTCACTGTTACCCCGTTGGATTGGCAGTGAAGTGTCTCAACCATGTCCAC 2113
 | || || |||| | || || ||||| ||||| || | || ||||

Db 2209 ATGGCTCCTGCCAGTGCTTTCTCTGGATGGATTGGCAAGGACTGCTCACAGGCTTGCCAC 2268

Qy 2114 CTGCCCCACTGGGGCCCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCA 2173
 | | ||||| |||| |||| | |||| |||| | || || ||||

Db 2269 CCGGTTCTGGGGCCCCGCTTCCACGCATGCAGCTGCCACAACGGGGCGAGCTGCA 2328

Qy 2174 GCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGA 2233
 |||| | || |||| || |||| ||||| || |||| |||| |||

Db 2329 GCGCCGAGGACGGGGCCTGCCACTGCACCCCTGGCTGGACTGGACTCTTCTGCACACAGC 2388

Qy 2234 GATGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACGGAG 2293
 | || || || |||| || || || || || || |||| |||| | || |

Db 2389 GCTGCCCAGCAGCATTTTTTGGGAAGGACTGTGGGCGCGTATGCCAGTGTGCAATGGCG 2448

Qy 2294 CTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTCTATGGGACGGCACT 2353
 | || ||||| || || |||| ||||| || |||| || | ||||

Db 2449 CCAGCTGTGACCACATCAGTGGCAAGTGCACCTGCCGCACAGGCTTCACGGGGCAACACT 2508

Qy 2354 GTGAGCAGAAGTGCCTTCAGGAACATATGGCTATGGCTGTGCGCAGATATGTGATTGTC 2413
 ||||| || | ||||| ||||| ||||| |||| |||| ||

Db 2509 GTGAGCAGAGATGTGCCCCAGGAACCTTTGGCTATGGGTGTGAGCAGCTATGTGAGTGCA 2568

Qy 2414 TGAACAACCTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGG 2473
 ||||| ||||| |||| || ||||| ||||| || || || ||

Db 2569 TGAACAACCTCCACCTGTGACCATGTCAACGGCACCTGTTACTGCAGCCCTGGCTTCAAAG 2628

Qy 2474 GAGCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAA 2533
 || || |||| ||||| || || || | | |||| || || |

Db 2629 GAATCAGGTGTGACCAAGCTGCCCTCATGATGGAGG---AGCTGAATCCCTACACCAAGA 2685

Qy 2534 CCAGTACTGCTCTCCCTGCTGATTCTTACCAGATCGGGGCCATTGCAGGCATCATATTC 2593
 ||| || || || || || || || || || || || || || ||

Db 2686 TCAGCCCAGCACTGGGTGCAGAGCGGCACTCGGTGGGTGCTGTACAGGCATCATGCTCC 2745

Qy 2594 TTGTCCTAGTTGTTCTCTTCTACTGGCATTGTTTCATTATTATAGACACAAGCAGAAGG 2653
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2746 TGTTATTCTCATTGTGGTGCTGCTGGGCCTATTGTCCTGGCATCGGCGGCGGCAGAAAG 2805
 Qy 2654 GAAAGGAATCAAGCATG---CCAGCAGTTACCTACACCCCTGCTATGAGGGTCGTCAATG 2710
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2806 AGAAGGGCCGAGACCTGGCTCCCCGTGTCTCCTACACACCTGCCATGAGGATGACCAGCA 2865
 Qy 2711 CAGATTATACCATTTCAGG 2729
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2866 CCGACTACTCCCTCTCAGG 2884

RESULT 2

AY406554

LOCUS AY406554 2910 bp DNA linear GSS 12-DEC-2003

DEFINITION Homo sapiens HCM2596 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY406554

VERSION AY406554.1 GI:39762525

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2910)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 2910)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..2910
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene <1..>2910
 /locus_tag="HCM2596"

ORIGIN

Query Match 29.0%; Score 992.6; DB 29; Length 2910;
 Best Local Similarity 57.8%; Pred. No. 2.3e-272;
 Matches 1396; Conservative 0; Mismatches 1013; Indels 6; Gaps 2;

Qy	318	CCCCCACTGTGCTGATAAAATGTGTCCATGGTCGCTGTATTGCTCCAAACACCTGTCAGTG	377
Db	75	CGCCCTGTGTACGGAGGAGTGTGTGCACGGCCGCTGCGTTTCCCGGACACCTGCCACTG	134
Qy	378	TGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGATGGTGATCACTGGGGTCC	437
Db	135	CGAGCCTGGCTGGGGAGGGCCCGACTGCTCCAGCGNNNNNNNNNNNNNNNNNNNNNNNNNN	194
Qy	438	CCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGCAACCCCATCACCGGGGC	497
Db	195	NN	254
Qy	498	TTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGCAGGGCAC	557
Db	255	NN	314
Qy	558	CTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCGACCACGT	617
Db	315	NN	374
Qy	618	CACGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTCTGTGAGGATCTTTGTCC	677
Db	375	NN	434
Qy	678	TCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAATGGAGGAGTGTG	737
Db	435	TCCTGGGAGCCATGGAGCTCACTGTGAGCTGCGCTGCCCTGTGAGAATGGGGGCACCTG	494
Qy	738	TCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGTGTGGTCA	797
Db	495	CCACCACATCACTGGCGAGTGTGCCTGCCCCCGAGGCTGGACGNNNNNNNNNNNNNNNNNN	554
Qy	798	GCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCCATAATGG	857
Db	555	NN	614
Qy	858	AGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAGGGGAACG	917
Db	615	NN	674
Qy	918	GTGCCAGGATGAGTGTCTGTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTG	977
Db	675	GTGCCAAGAGGAGTGCCCTTCGGGTCTTCGGCTTCCAGTGCTCACAGCGCTGTGACTG	734
Qy	978	TGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAGGCTTTGC	1037
Db	735	CCACAATGGGGGGCAGTGTTCACCCACCACGGGTGCCTGCGAGTGTGAGCCTGGCTACAA	794
Qy	1038	TGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGCTCTACGGCATCAAATGTGACAA	1097
Db	795	GGGCCACGCTGCCAGGAGCGACTGTGCCCGAGGGCCTGCATGGCCAGGCTGCACCCT	854
Qy	1098	ACGGTGTCCCTGCCACTTGAAAACACTCATAGCTGTACCCCATGTCTGGAGAGTGTGC	1157
Db	855	GCCCTGCCCTGTGACGCTGACAACACCATCAGCTGCCACCCAGTAAGTGGAGCTGTAC	914

Qy	1158	CTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGATTCTACGG	1217
Db	915	CTGCCAGCCAGGCTGGTCTGGTCACCACTGCAATGAATCCTGCCCTGTTGGCTACTATGG	974
Qy	1218	GGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACTGG	1277
Db	975	CGATGGCTGCCAGCTGCCTTGCACCTGTCAGAATGGCGCCGACTGCCACAGCATCACTGG	1034
Qy	1278	AAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTGGG	1337
Db	1035	GGGCTGCACTTGTGCTCCGGGCTTCATGGGAGAGGTCTGTGCCGTTTCTGTGCAGCAGG	1094
Qy	1338	AACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCTCC	1397
Db	1095	GACCTATGGCCCCAACTGCTCGTCCATCTGTAGCTGTAACAATGGTGGCACCTGCTCCCC	1154
Qy	1398	TGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGATG	1457
Db	1155	AGTAGATGGCTCCTGTACCTGCAAGGAAGGGTGGCAGGGCCTGGACTGCACCCTGCCATG	1214
Qy	1458	TCCCAGTGGCACATGGGGCTTTGGCTGTAACCTTAACATGCCAGTGCCTCAACGGGGGAGC	1517
Db	1215	TCCCAGTGGGACGTGGGGCTGAACTGCAACGAGAGCTGCACCTGTGCCAATGGGGCAGC	1274
Qy	1518	CTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGCGA	1577
Db	1275	CTGCAGCCCCATAGACGGCTCCTGCTCCTGCACTCCTGGCTGGCTGGGAGACACCTGTGA	1334
Qy	1578	ACTTCCCTGCCAGGATGGCACGTACGGGTGAACTGTGCTGAGCGCTGCGACTGCAGCCA	1637
Db	1335	GCTGCCTTGCCCGGATGGCACATTTGGGTGAACTGCAGTGAACACTGTGACTGCAGCCA	1394
Qy	1638	CGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGTGT	1697
Db	1395	TGCTGATGGATGTGACCCCGTCACAGGCCACTGCTGCTGCCTGGCCGGATGGACAGGCAT	1454
Qy	1698	CCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGCTA	1757
Db	1455	CCGCTGTGACAGCACGTGTCCACCTGGCCGCTGGGGCCCCAACTGCTCTGTCTCCTGCAG	1514
Qy	1758	CTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGCTT	1817
Db	1515	CTGTGAGAATGGAGGCTCCTGCTCCCCAGAGGATGGGAGCTGCGAGTGTGCCCTGGCTT	1574
Qy	1818	CCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTATGGGCATCGCTGCAGCCA	1877
Db	1575	CCGAGGACCCTTATGCCAGAGAACTGCCCCCTGGGTTCTATGGCCACGGCTGCGCCCA	1634
Qy	1878	GACATGCCCCACAGTGCCTTACAGCAGCGGGCCCTGCCACCACATCACGGGCCTGTGTGA	1937
Db	1635	GCCATGCCCCCTCTGCGTGCACAGCAGAGGCCCTGCCACCACATCAGCGGCATCTGTGA	1694
Qy	1938	CTGCTTGCCTGGCTTACAGGCGCCCTCTGCAATGAAGTGTGTCCAGTGGCAGATTTGG	1997
Db	1695	GTGCCTCCCAGGATTCTCTGGAGCTCTCTGCAACCAAGTGTGTGCTGGAGGATACTTTGG	1754
Qy	1998	GAAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTTGACAG	2057

Db	1755	GCAGGACTGTGCCAGCTCTGCTCCTGTGCCAACACGGGACCTGCAGCCCTATCGATGG	1814
Qy	2058	ATCTTGTCAGTGTTACCCCGGTTGGATTGGCAGTGACTGCTCTCAACCATGTCCACCTGC	2117
Db	1815	CTCCTGCCAGTGCTTTCTGGATGGATTGGCAAGGACTGCTCACAGGCTTGCCACCCGG	1874
Qy	2118	CCACTGGGGCCCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGCGC	2177
Db	1875	GTTCTGGGGCCCCGCTGCTTCCACGCATGCAGCTGCCACAACGGGGCGAGCTGCAGCGC	1934
Qy	2178	CTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGATG	2237
Db	1935	CGAGGACGGGGCCTGCCACTGCACCCCTGGCTGGACTGGACTCTTCTGCACACAGCGCTG	1994
Qy	2238	TCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCTGA	2297
Db	1995	CCCAGCAGCATTTTTTGGGAAGGACTGTGGGCGCGTATGCCAGTGTGAGAATGGCGCCAG	2054
Qy	2298	CTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTCATGGGACGGCACTGTGA	2357
Db	2055	CTGTGACCACATCAGTGGCAAGTGACCTGCCGCACAGGCTTCACCGGGCAACACTGTGA	2114
Qy	2358	GCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGTCTGAA	2417
Db	2115	GCAGAGATGTGCCCCAGGAACCTTTGGCTATGGGTGTCAGCAGCTATGTGAGTGCATGAA	2174
Qy	2418	CAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGAGC	2477
Db	2175	CAACTCCACCTGTGACCATGTACCGGCACCTGTTACTGCAGCCCTGGCTTCAAAGGAAT	2234
Qy	2478	GAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACCAG	2537
Db	2235	CAGGTGTGACCAAGCTGCCCTCATGATGGAGG---AGCTGAATCCCTACACCAAGATCAG	2291
Qy	2538	TACTGCTCTCCCTGCTGATTCCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTTGT	2597
Db	2292	CCCAGCACTGGGTGCAGAGCGGCACTCGGTGGGTGCTGTACAGGCATCATGCTCCTGTT	2351
Qy	2598	CCTAGTTGTTCTCTTCTACTGGCATTGTTCAATTATTTATAGACACAAGCAGAAGGGAAA	2657
Db	2352	ATTCTCATTTGTGGTGTCTGCTGGGCCTATTTGCCTGGCATCGGCGGCGGCAGAAAGAGAA	2411
Qy	2658	GGAATCAAGCATG---CCAGCAGTTACCTACACCCCTGCTATGAGGGTCGTCAATGCAGA	2714
Db	2412	GGGCCGAGACCTGGCTCCCCGTGTCTCTACACACCTGCCATGAGGATGACCAGACCGA	2471
Qy	2715	TTATACCATTTCAGG	2729
Db	2472	CTACTCCCTCTCAGG	2486

RESULT 3

AK051642

LOCUS AK051642 3568 bp mRNA linear HTC 20-SEP-2003

DEFINITION Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone:D130061K05 product:MEGF11 PROTEIN

(KIAA1781) homolog [Homo sapiens], full insert sequence.

ACCESSION AK051642

VERSION AK051642.1 GI:26342079

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3568)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,

Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers

source 1. .3568
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:D130061K05"
/db_xref="MGI:2419956"
/db_xref="taxon:10090"
/clone="D130061K05"
/tissue_type="spinal ganglion"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"

CDS 147. .2990
/note="unnamed protein product; MEGF11 PROTEIN (KIAA1781)
homolog [Homo sapiens] (SPTR|Q96KG6, evidence: FASTY,
85.9%ID, 85.4%length, match=2363)
putative"
/codon_start=1
/protein_id="BAC34702.1"
/db_xref="GI:26342080"
/translation="MAPSAVGLLVFLLQAALALNPEDPNVCSHWESYAVTVQESYAHF
FDQIYYTRCADILNWFKCTRHRISYKTAYRRGLRTMYRRRSQCCPGYYENGDFCIRCD
SEHWGPHCSNRCQCQNGALCNPIITGACVCAPGFRGWRCEELCAPGTHGKGCQLLCQCH
HGASCDPRTGECCLCAPGYTGVCCEELCPPGSHGAHCELRCPQNGGTCHHITGECACP
PGWTGAVCAQPCPPGTFGQNCSDCPCHHGGQCDHVTGQCHCTAGYMGDRCEECFPG
TFGFLCSQRCDCHNGGQCSPATGACECEPGYKGPSCQERLCPEGLHGPCTLPCCPDT
ENTISCHPVTGACTCQPGWSGHYCNESCPAGYYGNGCQLPCTCQNGADCHSITGSCTC
APGFMGEVCAVPCAAGTYGPNCSVCSNNGGTCSPVDGSCTCREGWQGLDCSLPCPS
GTWGLNCNETCICANGAACSPFDGSCACTPGWLGDSCLEPCPDGTFGLNCSEHCDCSH
ADGCDPVTGHCCCLAGWTGIRCDSTCPPGRWGPNCVSCSCENGGSCEPDGSCCAP
GFRGPLCQRICPPGFYGHGCAQPCPLCVHSRGPCHHISGICECLPGFSGALCNQVCAG
GHFGQDCAQLCSCANNGTCSPIDGSCQCFPGWIGKDCSQGCPSAFFGKDCGHICQCQN
GASCDHITGKCTCRTGFSGRHCEQRCAPGTFGYGCQQLCECMNNATCDHVTGTCTCYCSP
GFKGIRCDQAALMDELNPYTKISPALGAERHSVGAVTGIVLLLFLVVLLGLFAWRR
RRQKEKGRDLAPRVSYTPAMRMTSTDYSLSDLSSSSHAQCFSNASYHTLACGGPATS
QASTLDRNSPTKLSNKS LDRDTAGWTPYSYVNVLD SHFQISALEARYPPEDFYIELRH
LSRHAEPHSPGTCGMDRRQNTYIMDKGFKVAPA"

ORIGIN

Query Match 28.4%; Score 972.6; DB 11; Length 3568;
Best Local Similarity 65.4%; Pred. No. 1.4e-266;
Matches 1495; Conservative 0; Mismatches 699; Indels 93; Gaps 1;

```
Qy      74 CTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACTCAGTGACTGTGC 133
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      199 CTCTGAACCCTGAAGACCCCAATGTGTGTAGCCACTGGGAGAGCTATGCCGTGACTGTGC 258

Qy      134 AAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGCACTGACATTCTAA 193
      | |||| | |||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      259 AGGAGTCTTATGCACACCCCTTTGATCAGATCTACTACACAGATGTGCAGACATCCTCA 318

Qy      194 ACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTATCGACATGGGGAGA 253
      ||||| || || || ||||| | ||||| || || || || || ||
Db      319 ACTGGTTCAAGTGTACCGGCGACCGGATCAGCTATAAGACCGCGTATAGGCGCGGCTCC 378

Qy      254 AGACTATGTATAGGCGCAAGTCTCAGTGTGTCTGGATTTTATGAAAGCGGGGAAATGT 313
      ||| |||| |||| | || || || || || |||| | |||| | ||| |
Db      379 GGACCATGTACCGGCGGAGGTCCCAATGCTGCCCTGGCTACTATGAGAACGGAG----- 432

Qy      314 GTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCTCCAAACACCTGTC 373
Db      433 ----- 432

Qy      374 AGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGATGGTGATCACTGGG 433
                        ||| || || | ||| || | || |||||
Db      433 -----ACTTCTGCATTCGCTGTGACAGCGAGCACTGGG 465

Qy      434 GTCCCCACTGCACCAGCCGGTGCCAGTGCAAAATGGGGCTCTGTGCAACCCCATCACCG 493
      ||||| ||||| || ||||| ||||| || || || ||||| ||||| |||||
Db      466 GTCCCCACTGCAGCAACCGGTGTCAGTGTGAGAACGGCGCCCTGTGCAACCCATCACCG 525

Qy      494 GGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGCAGG 553
      | || ||| || || ||||| ||||| ||||| ||||| || || | | |
Db      526 GCGCCTGCGTGTGCGCCCCGGGCTTCCGAGGCTGGCGCTGTGAGGAACTCTGCGCTCCTG 585

Qy      554 GCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCGACC 613
      | || | || || || || || || || ||||| |||| |||| || ||||
Db      586 GTA CTCACGGCAAGGGCTGCCAGCTGCTCTGTGAGTGCCACCATGGCGCCAGCTGTGACC 645

Qy      614 ACGTCACGGGGGAATGCCGCTGCCACCAGGATACACGGAGCCTTCTGTGAGGATCTTT 673
      ||| || || |||| |||| | || || ||||| || | | ||||| || ||
Db      646 CGCGCACTGGCGAGTGCCTCTGCGCTCCTGGCTACACAGGCGTTACTGTGAGGAGCTGT 705

Qy      674 GTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAAATGGAGGAG 733
      | || ||||| | |||| | || ||||| | ||||| || || |||||
Db      706 GCCCCCTGGGAGCCATGGAGCTACTGTGAGCTGCGCTGCCCCCTGCCAGAATGGAGGCA 765

Qy      734 TGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGTGTG 793
      || || || ||||| |||| | ||||| | ||||| || |||||
Db      766 CCTGCCACCACATCACTGGCGAATGTGCCTGCCCTCCAGGCTGGACGGGAGCAGTGTGTG 825

Qy      794 GTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCCATA 853
      ||||| ||||| || ||||| ||||| || || || || || |||||
Db      826 CCCAGCCCTGCCCTCCAGGGACCTTTGGCCAGAAGTGTAGCCAGGACTGTCCCTGCCACC 885

Qy      854 ATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAGGGG 913
```

Db	886	ATGGAGGCCAGTGTGACCATGTGACTGGACAATGCCACTGTACAGCTGGATACATGGGGG	945
Qy	914	AACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGCC	973
Db	946	ACAGGTGTCAAGAAGAATGTCCCTTTGGAACGTTTCGGTTTCCTGTGCTCTCAACGCTGTG	1005
Qy	974	AGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCCTCTGTGAAGCAGGCT	1033
Db	1006	ACTGCCACAATGGAGGTCAATGTTTACCAGCCACAGGGGCTGTGAGTGTGAGCCTGGCT	1065
Qy	1034	TTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGCTCTACGGCATCAAATGTG	1093
Db	1066	ACAAGGGCCCTAGCTGCCAGGAGCGGCTATGCCCTGAGGGCTGCATGGCCCAGGCTGCA	1125
Qy	1094	ACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTCACCCCATGTCTGGAGAGT	1153
Db	1126	CCTTGCCCTGCCCCCTGTGACACCGAGAACACTATCAGCTGCCATCCAGTTACTGGAGCTT	1185
Qy	1154	GTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCTCGGATTCT	1213
Db	1186	GTACCTGCCAACCAGGCTGGTCTGGCCACTACTGCAATGAGTCCTGCCCCGCCGGCTACT	1245
Qy	1214	ACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGA	1273
Db	1246	ATGGCAACGGTTGCCAGCTACCCTGCACCTGCCAGAACGGTGCTGACTGCCACAGTATCA	1305
Qy	1274	CTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTC	1333
Db	1306	CCGGGAGCTGCACTTGTGCTCCAGGCTTCATGGGAGAGGTGTGTGCCGTCCCCTGTGCTG	1365
Qy	1334	TGGGAACCTATGGGATAAACTGTTTCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCT	1393
Db	1366	CAGGGACCTATGGTCCCAACTGTTTCTGTATGTAGCTGTAGCAACGGCGGCACCTGTT	1425
Qy	1394	CTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCA	1453
Db	1426	CCCCAGTGGATGGCTCCTGCACCTGCCGAGAGGGATGGCAGGGCCTGGACTGCTCCCTGC	1485
Qy	1454	GATGTCCCAGTGGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGG	1513
Db	1486	CTTGTCCCAGTGGGACCTGGGGCCTGAACTGCAATGAGACTTGCATCTGTGCCAATGGAG	1545
Qy	1514	GAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAAT	1573
Db	1546	CTGCCTGCAGCCCCTTTGATGGGTCCTGTGCCTGCACCCAGGCTGGCTGGGGGACTCCT	1605
Qy	1574	GCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCA	1633
Db	1606	GTGAAGTGGCCTGCCCGGACGGCACTTTTGGGCTGAACTGCAGTGAGCATTGCGACTGCA	1665
Qy	1634	GCCACGCAGATGGCTGCCACCCTACCACGGGGCCATTGCCGCTGCCTCCCGGGATGGTCAG	1693
Db	1666	GCCATGCTGATGGCTGTGACCCTGTACAGGCCACTGCTGCTGCCTGGCAGGATGGACAG	1725
Qy	1694	GTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCT	1753

Db 1726 GCATCCGCTGTGATAGCACGTGTCTCCAGGTCGCTGGGGCCCCAACTGTTTCAGTGTCTCT 1785
 Qy 1754 GCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAG 1813
 || |||| | || | | ||| ||||| || || | ||||| ||| |
 Db 1786 GCAGCTGTGAGAACGGAGGTTCTGTCTCCCGGAGGACGGGAGCTGCGAGTGTGCCCTG 1845
 Qy 1814 GCTTCCGAGGCACCACTTGTCTCAGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCTGCA 1873
 |||| |||| | | ||||| ||||| | || | | || | ||||
 Db 1846 GCTTTTCGAGGACCCTTATGTCTCAGAGAATCTGCCACCAGGATTCTACGGCCATGGCTGCG 1905
 Qy 1874 GCCAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGT 1933
 |||| | || | | |||| ||||| ||||| ||||| || | |
 Db 1906 CCCAGCCTTGTCCCCTCTGCGTGCACAGCAGGGGGCCCTGCCACCACATCAGTGGTATCT 1965
 Qy 1934 GTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGAT 1993
 |||| || |||| | || | | || | |||| ||||| | | || |
 Db 1966 GTGAGTGCCTGCCAGGATTCTCTGGAGCCTTGTGCAACCAAGTGTGTGCTGGAGGGCACT 2025
 Qy 1994 TTGGGAAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTTG 2053
 | || | ||||| | || | || | |||| | |||| | ||||| |
 Db 2026 TCGGGCAGGACTGTGCCAGCTCTGTTCTGTGCCAACCAATGGGACCTGCAGCCCCATCG 2085
 Qy 2054 ACAGATCTTGTCTAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTCTCAACCATGTCCAC 2113
 | | || ||||| | || | ||||| ||||| || | || |
 Db 2086 ATGGCTCCTGTCTAGTGTCTCCCTGGGTGGATTGGCAAGGACTGTCTCAGGGGTGCCCTT 2145
 Qy 2114 CTGCCCCACTGGGGCCCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCA 2173
 | || | || || |||| |||| || | || | ||||| |||
 Db 2146 CAGCATTTTTTGGGAAGGACTGTGGGCACATATGCCAGTGTCTCAGAATGGAGCCAGCTGTG 2205
 Qy 2174 GCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGA 2233
 | | | |||| |||| | || | |||| | || | |||| ||||
 Db 2206 ACCACATCACTGGGAAATGCACCTGTCTGAACAGGCTTCTCTGGCCGCCACTGTGAACAGA 2265
 Qy 2234 GATGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACGGAG 2293
 |||| | | || | |||| | | || | | |||| | || ||| |
 Db 2266 GATGTGCCCTGGAACCTTTGGATATGGGTGTCTCAGCAGCTATGTGAGTGCATGAACAATG 2325
 Qy 2294 CTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGCACT 2353
 | || ||||| | |||| || | || | ||||| ||| |
 Db 2326 CCACTTGTGACCACGTCACTGGTACCTGTTACTGTAGCCCGGGATTCAAAGGAATCAGGT 2385
 Qy 2354 GTGAGCA 2360
 |||| ||
 Db 2386 GTGACCA 2392

RESULT 4

AY406556

LOCUS AY406556 2910 bp DNA linear GSS 12-DEC-2003

DEFINITION Mus musculus HCM2596 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY406556

VERSION AY406556.1 GI:39762527

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Qy 678 TCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAAATGGAGGAGTGTG 737
 ||||| || | | | | | | | | | | | | | | | |
 Db 435 TCCTGGTACTCACGGCAAGGGCTGCCAGCTGCTCTGTCAAGTCCACCATGGCGCCAGCTG 494

Qy 738 TCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGTGTGGTCA 797
 | | | | | | | | | | | | | | | |
 Db 495 TGACCCGCGCACTGGCGAGTGCCTCTGCGCTCCTGGCTACACANNNNNNNNNNNNNNNNN 554

Qy 798 GCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCCATAATGG 857
 Db 555 NNN 614

Qy 858 AGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAGGGGAACG 917
 Db 615 NNN 674

Qy 918 GTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTG 977
 ||| || | | | | | | | | | | | | | | | | | | | | |
 Db 675 GTGTCAAGAAGAATGTCCCTTTGGAACGTTTCGGTTTCTGTGCTCTCAACGCTGTGACTG 734

Qy 978 TGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAGGCTTTGC 1037
 ||| |||| | |||| | | | | | | | | | | | | | | |
 Db 735 CCACAATGGAGGTCAATGTTACCAGCCACAGGGGCCTGTGAGTGTGAGCCTGGCTACAA 794

Qy 1038 TGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTACGGCATCAAATGTGACAA 1097
 ||| |||| | | | | | | | | | | | | | | | | | |
 Db 795 GGGCCCTAGCTGCCAGGAGCGGCTATGCCCTGAGGGCCTGCATGGCCAGGCTGCACCTT 854

Qy 1098 ACGGTGTCCCTGCCACTTGGAACACTCATAGCTGTCACCCCATGTCTGGAGAGTGTGC 1157
 | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 855 GCCCTGCCCCCTGTGACACCGAGAACACTATCAGCTGCCATCCAGTTACTGGAGCTTGTA 914

Qy 1158 CTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGATTCTACGG 1217
 |||| | | | | | | | | | | | | | | | | | | | | | |
 Db 915 CTGCCAACCAGGCTGGTCTGGCCACTACTGCAATGAGTCCTGCCCCGCCGGCTACTATGG 974

Qy 1218 GGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACTGG 1277
 | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 975 CAACGGTTGCCAGCTACCCTGCACCTGCCAGAACGGTGTGACTGCCACAGTATCACCGG 1034

Qy 1278 AAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTGGG 1337
 | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1035 GAGCTGCACTTGTGCTCCAGGCTTCATGGGAGAGGTGTGTGCCGTCCCCTGTGCTGCAGG 1094

Qy 1338 AACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCTCC 1397
 ||||| || | | | | | | | | | | | | | | | | | | | | |
 Db 1095 GACCTATGGTCCCAACTGTTTCATCTGTATGTAGCTGTAGCAACGGCGGCACCTGTTCCCC 1154

Qy 1398 TGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGATG 1457
 |||| | | | | | | | | | | | | | | | | | | | | | |
 Db 1155 AGTGGATGGCTCCTGCACCTGCCGAGAGGGATGGCAGGGCCTGGACTGCTCCCTGCCCTTG 1214

Qy 1458 TCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCCTAACGGGGGAGC 1517
 ||||| || | | | | | | | | | | | | | | | | | | | | |
 Db 1215 TCCCAGTGGGACCTGGGGCCTGAACTGCAATGAGACTTGCATCTGTGCCAATGGAGCTGC 1274

Qy 1518 CTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGCGA 1577
 |||| | | | | | | | | | | | | | | | | | | | |
 Db 1275 CTGCAGCCCTTTGATGGGTCTGTGCCTGCACCCAGGCTGGCTGGGGGACTCCTGTGA 1334

Qy 1578 ACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGCCA 1637
 ||| ||||| ||| |||| | | ||||| |||| | ||||| |||||
 Db 1335 ACTGCCCTGCCCGGACGGCACTTTTGGGCTGAACTGCAGTGAGCATTGCGACTGCAGCCA 1394

Qy 1638 CGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGTGT 1697
 || ||||| |||| | | |||| | | |||| | | |||| | | |
 Db 1395 TGCTGATGGCTGTGACCCTGTCACAGGCCACTGCTGCTGCCTGGCAGGATGGACAGGCAT 1454

Qy 1698 CCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGCTA 1757
 || ||||| ||| ||| || | | ||||| |||| | | ||||
 Db 1455 CCGCTGTGATAGCACGTGTCTCCAGGTCGTGGGGCCCCAACTGTTCACTGTCTGTCAG 1514

Qy 1758 CTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGCTT 1817
 |||| | | | | | | ||||| | | | | | | ||||| ||| |||||
 Db 1515 CTGTGAGAACGGAGGTTCTGCTCCCCGGAGGACGGGAGCTGCGAGTGTGCCCTGGCTT 1574

Qy 1818 CCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTATGGGCATCGCTGCAGCCA 1877
 |||| | | ||||| |||| | | | | | | | | |||| | |
 Db 1575 TCGAGGACCCTTATGTGAGAGAATCTGCCACCAGGATTCTACGGCCATGGCTGCGCCCA 1634

Qy 1878 GACATGCCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGTGA 1937
 | | | | | | |||| ||||| ||||| ||||| || | ||||
 Db 1635 GCCTTGTCCCCTCTGCGTGCACAGCAGGGGGCCCTGCCACCACATCAGTGGTATCTGTGA 1694

Qy 1938 CTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTGG 1997
 || |||| | | || | | || | | |||| | | | | ||||
 Db 1695 GTGCCTGCCAGGATTCTCTGGAGCCTTGTGCAACCAAGTGTGTGCTGGAGGGCACTTCGG 1754

Qy 1998 GAAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTTGACAG 2057
 | | ||||| | ||| ||| ||||| | |||| | ||||| || |
 Db 1755 GCAGGACTGTGCCAGCTCTGTTCTGTGCCAACAATGGGACCTGCAGCCCCATCGATGG 1814

Qy 2058 ATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGACTGCTCTCAACCATGTCCACCTGC 2117
 || ||||| || | || | | ||||| || ||||| || | || ||||
 Db 1815 CTCCTGTGAGTGTCTCCCTGGGTGGATTGGCAAGGACTGCTCACAGGCCTGCCCATCTGG 1874

Qy 2118 CCACTGGGGCCCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTCTGCAGCGC 2177
 ||||| | ||| ||||| ||| |||| | | || | |||||
 Db 1875 GTTCTGGGGCTCTGCCTGCTTCCACACATGCAGCTGCCACAACGGGGCGAGCTGCAGCGC 1934

Qy 2178 CTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGATG 2237
 | | ||||| || | |||| ||||| || | ||| ||||| || | ||
 Db 1935 CGAGGATGGGGCCTGCCACTGCACCCCTGGCTGGACTGGACTCTTCTGCACGCAGCGTTG 1994

Qy 2238 TCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCTGA 2297
 ||| || |||| ||| || | || | | ||||| |||| | |||||
 Db 1995 CCCTTCAGCATTTTTTGGGAAGGACTGTGGGCACATATGCCAGTGTGAGAATGGAGCCAG 2054

Qy 2298 CTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGCACTGTGA 2357
 ||| ||||| |||| | | | | | | | | | | || || |||||
 Db 2055 CTGTGACCACATCACTGGGAAATGCACCTGTGAACAGGCTTCTCTGGCCGCCACTGTGA 2114

Qy 2358 GCAGAAAGTCCCTTCAGGAACATATGGCTATGGCTGTCGCCAGATATGTGATTGTCTGAA 2417

Db	2115	ACAGAGATGTGCCCTGGAACCTTTGGATATGGGTGTCAGCAGCTATGTGAGTGCATGAA	2174
Qy	2418	CAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGAGC	2477
Db	2175	CAATGCCACTTGTGACCACGTCACCTGGTACCTGTTACTGTAGCCCGGGATTCAAAGGAAT	2234
Qy	2478	GAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACCAG	2537
Db	2235	CAGGTGTGACCAAGCTGCCCTCATGATGGATG---AGCTGAATCCCTACACCAAGATCAG	2291
Qy	2538	TACTGCTCTCCCTGCTGATTCTTACCAGATCGGGGCCATTGCAGGCATCATCATTTCTTGT	2597
Db	2292	TCCAGCTCTGGGAGCAGAGCGGCACTCAGTGGGTGCTGTCACCGGCATCGTTCTCCTGTT	2351
Qy	2598	CCTAGTTGTTCTCTTCTACTGGCATTGTTTCATTATTTATAGACACAAGCAGAAGGGAAA	2657
Db	2352	GTTCTTGGTGGTGGTGGTCTGCTGGGCCCTGTTTGCCCTGGCGACGGAGGCGGCAGAAAGAGAA	2411
Qy	2658	GGAATCAAGCATG---CCAGCAGTTACCTACACCCCTGCTATGAGGGTCGTCAATGCAGA	2714
Db	2412	AGGCCGTGACCTGGCTCCCCGAGTCTCTACACCCAGCCATGAGGATGACCAGCACAGA	2471
Qy	2715	TTATACCATTTTCAGGAAC	2732
Db	2472	CTACTCTCTCTCAGGCAC	2489

AK032661

LOCUS	AK032661	3466 bp	mRNA	linear	HTC 18-SEP-2003
-------	----------	---------	------	--------	-----------------

DEFINITION Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length

(KIAA1781) homolog [Homo sapiens], full insert sequence.

VERSION AK032661.1 GI:26082960

KEYWORDS HTC; CAP trapper.

SOURCE Has musculus (house mouse),
ORGANISM Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3466)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL:http://genome.gsc.riken.go.jp/

URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers

source 1. .3466

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:6530404N23"


```
misc_feature 1. .3466
/db_xref="MGI:2396257"
/db_xref="taxon:10090"
/clone="6530404N23"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
/note="MEGF11 PROTEIN (KIAA1781) homolog [Homo sapiens]
(SPTR|Q96KG6, evidence: FASTY, 85.9%ID, 85.4%length,
match=2363)"
```

ORIGIN

```
Query Match          26.1%;  Score 892.6;  DB 11;  Length 3466;
Best Local Similarity 69.1%;  Pred. No. 1.3e-243;
Matches 1219;  Conservative 0;  Mismatches 544;  Indels 0;  Gaps 0;
```

```
Qy      94 AATGTGTGTAGCCACTGGGAAAGCTACTCAGTGACTGTGCAAGAGTCATACCCACATCCC 153
      |||
Db      1 AATGTGTGTAGCCACTGGGAGAGCTATGCCGTGACTGTGCATGAGTCTTATGCACACCCC 60

Qy     154 TTTGATCAAATTTACTACACGAGCTGCACTGACATTCTAAACTGGTTTAAATGCACGCGG 213
      |||
Db      61 TTTGATCAGATCTACTACACACGATGTGCAGACATCCTCAACTGGTTCAGTGTACCCGG 120

Qy     214 CACAGAGTCAGCTATCGGACAGCCTATCGACATGGGGAGAAGACTATGTATAGGCGCAAG 273
      |||
Db     121 CACCGGATCAGCTATAAGACCGCGTATAGGCGCGGCCTCCGGACCATGTACCGGCGGAGG 180

Qy     274 TCTCAGTGTGTCTGCTGGATTTTATGAAAGCGGGGAAATGTGTGTCCCCCACTGTGCTGAT 333
      |||
Db     181 TCCAATGCTGCCCTGGCTACTATGAGAACGGAGACTTCTGCATTCCTCTGTGTACCGAG 240

Qy     334 AAATGTGTCCATGGTTCGCTGTATTGCTCCAAACACCTGTCAGTGTGAGCCTGGCTGGGGA 393
      |||
Db     241 GAGTGCATGCACGGCCGCTGTGTCTCTCCGATACCTGCCACTGTGAGCCTGGATGGGGA 300

Qy     394 GGGACCAACTGCTCCAGTGCCTGCGATGGTGATCACTGGGGTCCCCACTGCACCAGCCGG 453
      |||
Db     301 GGCCCTGACTGCTCCAGCGGCTGTGACAGCGAGCACTGGGGTCCCCACTGCAGCAACCGG 360

Qy     454 TGCCAGTGCAAAAATGGGGCTCTGTGCAACCCCATCACCGGGGCTTGCCACTGTGCTGCG 513
      |||
Db     361 TGTCAGTGTGACAACGGCGCCCTGTGCAACCCTATCACCGGCGCCTGCGTGTGCGCCCCG 420

Qy     514 GGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGCAGGGCACCTATGGTAACGACTGT 573
      |||
Db     421 GGCTTCCGAGGCTGGCGCTGTGAGGAACCTCTGCGCTCCTGGTACTCACGGCAAGGGCTGC 480

Qy     574 CATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCGACCACGTACGGGGGAATGCCGC 633
      |||
Db     481 CAGCTGCTCTGTCACTGCCACCATGGCGCCAGCTGTGACCCGCGCACTGGCGAGTGCCTC 540

Qy     634 TGCCCACCAGGATACACCGGAGCCTTCTGTGAGGATCTTTGTCTCCTGGTAAACATGGT 693
      |||
Db     541 TGCGCTCCTGGCTACACAGGCGTTTACTGTGAGGAGCTGTGCCCCCTGGGAGCCATGGA 600

Qy     694 CCACAGTGTGAGCAGAGATGCCCTTGTCAAATGGAGGAGTGTGTCATCACGTCACTGGA 753
```

Db	601	GCTCACTGTGAGCTGCGCTGCCCTGCCAGAAATGGAGGCACCTGCCACCACATCACTGGC	660
Qy	754	GAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGT	813
Db	661	GAATGTGCCTGCCCTCCAGGCTGGACGGGAGCAGTGTGTGCCAGCCCTGCCCTCCAGGG	720
Qy	814	CGCTTTGGAAAGAACTGTTCCCAAGAAATGCCAGTGCCATAATGGAGGGACGTGTGATGCT	873
Db	721	ACCTTTGGCCAGAAGTGTAGCCAGGACTGTCCCTGCCACCATGGAGGCCAGTGTGACCAT	780
Qy	874	GCCACAGGCCAATGTCATTGCAGTCCAGGATACACAGGGGAACGGTGCCAGGATGAGTGT	933
Db	781	GTGACTGGACAATGCCACTGTACAGCTGGATACATGGGGGACAGGTGTCAAGAAGAATGT	840
Qy	934	CCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAG	993
Db	841	CCCTTTGGAACGTTTCGTTTCTGTGCTCTCAACGCTGTGACTGCCACAATGGAGGTCAA	900
Qy	994	TGTTACCACGTGAGCGGCGCATGCCCTCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCCAA	1053
Db	901	TGTTACCAGCCACAGGGGCTGTGAGTGTGAGCCTGGCTACAAGGGCCCTAGCTGCCAG	960
Qy	1054	GCACGCCTGTGTCTGAGGGGCTCTACGGCATCAAATGTGACAAACGGTGTCCCTGCCAC	1113
Db	961	GAGCGGCTATGCCCTGAGGGCTGCATGGCCAGGCTGCACCTTGCCCTGCCCTGTGAC	1020
Qy	1114	TTGGAAAACACTCATAGCTGTCACCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGG	1173
Db	1021	ACCGAGAACACTATCAGCTGCCATCCAGTTACTGGAGCTTGTACCTGCCAACCAGGCTGG	1080
Qy	1174	TCAGGACTCTACTGTAATGAGACATGTTCTCTGGATTCTACGGGGAAGCTTGCCAGCAG	1233
Db	1081	TCTGGCCACTACTGCAATGAGTCTGCCCGCCGGCTACTATGGCAACGGTTGCCAGCTA	1140
Qy	1234	ATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCC	1293
Db	1141	CCCTGCACCTGCCAGAACGGTGTGACTGCCACAGTATCACCGGGAGCTGCACTTGTGCT	1200
Qy	1294	CCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAAC	1353
Db	1201	CCAGGCTTCATGGGAGAGGTGTGTGCCGTCCCCTGTGCTGCAGGGACCTATGGTCCCAAC	1260
Qy	1354	TGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGT	1413
Db	1261	TGTTCACTCTGTATGTAGCTGTAGCAACGGCGGCACCTGTCCCCAGTGGATGGCTCCTGC	1320
Qy	1414	ACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGATGTCCCAGTGGCACATGG	1473
Db	1321	ACCTGCCGAGAGGGATGGCAGGGCCTGGACTGCTCCCTGCCTTGTCCCAGTGGGACCTGG	1380
Qy	1474	GGCTTTGGCTGTAACTTAACATGCCAGTGCCTCAACGGGGAGCCTGCAACACCCTGGAC	1533
Db	1381	GGCTGAACTGCAATGAGACTTGCATCTGTGCCAATGGAGCTGCCTGCAGCCCCCTTGTAT	1440
Qy	1534	GGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGAT	1593

```

Db      1441 GGGTCCTGTGCCTGCACCCCAGGCTGGCTGGGGGACTCCTGTGAACTGCCCTGCCCCGGAC 1500
Qy      1594 GGCACGTACGGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCAC 1653
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1501 GGCACCTTTTGGGCTGAACTGCAGTGAGCATTGCGACTGCAGCCATGCTGATGGCTGTGAC 1560
Qy      1654 CCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTG 1713
      ||| ||| ||||| ||| ||||| | ||||| ||||| ||| ||||| ||| |
Db      1561 CCTGTACAGGCCACTGCTGCTGCCTGGCAGGATGGACAGGCATCCGCTGTGATAGCACG 1620
Qy      1714 TGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCT 1773
      ||| || | ||||| ||||| || || ||||| ||||| || || || ||
Db      1621 TGTCTCCAGGTCGCTGGGGCCCCAACTGTTTCAGTGTCTGCAGCTGTGAGAACGGAGGT 1680
Qy      1774 TCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGCTTCCGAGGCACCACCTTGT 1833
      || ||||| || || || || ||||| || ||||| ||||| || || ||
Db      1681 TCCTGCTCCCCGGAGGACGGGAGCTGCGAGTGTGCCCTGGCTTTCGAGGACCCTTATGT 1740
Qy      1834 CAGAGGATCTGCTCCCCTGGTTT 1856
      ||||| | | |||| |||
Db      1741 CAGAGAAGTAAGGCTCCTGATTT 1763

```

RESULT 6

AY406555

LOCUS AY406555 2910 bp DNA linear GSS 12-DEC-2003
 DEFINITION Pan troglodytes HCM2596 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY406555

VERSION AY406555.1 GI:39762526

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1 (bases 1 to 2910)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 2910)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1. .2910

/organism="Pan troglodytes"

[illegible]

Qy	1038	TGGCGAGCGCTGCGAAGCACGCCTGTGTCCCTGAGGGGCTCTACGGCATCAAATGTGACAA	1097
Db	795	NN	854
Qy	1098	ACGGTGTCCCTGCCACTTGGAACACTCATAGCTGTCACCCCATGTCTGGAGAGTGTGC	1157
Db	855	NN	914
Qy	1158	CTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGATTCTACGG	1217
Db	915	NN	974
Qy	1218	GGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACTGG	1277
Db	975	NN	1034
Qy	1278	AAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTGGG	1337
Db	1035	NN	1094
Qy	1338	AACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCTCC	1397
Db	1095	NN	1154
Qy	1398	TGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGATG	1457
Db	1155	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGGTGGCAGGGCCTGGACTGCACCCTGCCATG	1214
Qy	1458	TCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGAGC	1517
Db	1215	TCCCAGTGGGACATGGGGCCTGAACTGCAACGAGAGCTGCACCTGTGCCAATGGGGCAGC	1274
Qy	1518	CTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGCGA	1577
Db	1275	CTGNNGCCCCATNGNCGGCTCCTGCTCCTGCACTCCTGGCTGGCTGGGAGNCACCTGTGA	1334
Qy	1578	ACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGCCA	1637
Db	1335	GCTGCCCTGCCCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGCAGTGAACACTGTGACTGCAGCCN	1394
Qy	1638	CGCAGATGGCTGCCACCCTACCACGGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGTGT	1697
Db	1395	NNNNNNNTGGATGTGACCCCGTCACAGGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGCAT	1454
Qy	1698	CCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGCTA	1757
Db	1455	CCACTGTGACAGCACGTGTCCACCTGGCCGCTGGGGCCCCAACTGCTCTGTCTCCTGCAG	1514
Qy	1758	CTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGCTT	1817
Db	1515	CTGTGAGAATGGAGGCTCCTGCTCCCGAGAGGATGGGAGCTGCGAGTGTGCCCTGGCTT	1574
Qy	1818	CCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCTGCAGCCA	1877
Db	1575	CCGAGGACCCTTATGCCAGAGAATCTGCCCCCTGGGTTCTATGGCCACGGCTGCGCCCA	1634

[illegible]

Db || |
2472 CTACNNNNNNNNNG 2486

RESULT 7

AK053551

LOCUS AK053551 3556 bp mRNA linear HTC 20-SEP-2003

DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:E130107B17 product:similar to MEGF12 [Mus musculus], full insert sequence.

ACCESSION AK053551

VERSION AK053551.1 GI:26343538

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,

Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers

source 1. .3556
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:E130107B17"
/db_xref="MGI:2425343"
/db_xref="taxon:10090"
/clone="E130107B17"
/tissue_type="eyeball"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
CDS 205. .3309

/note="unnamed protein product; putative similar to MEGF12 [Mus musculus] (SPTR|AAL33583, evidence: FASTY, 72.7%ID, 41.5%length, match=1604)"
/codon_start=1
/protein_id="BAC35426.1"
/db_xref="GI:26343539"
/translation="MPLCPLLLLALGLRLTGTLSNDPNVCTFWESFTTTTKEHLRP
FSLIPAESCHRPWEDPHTCAQPTVVYRTVYRQVVKMDSRPRLQCCRGYYESRGACVPL
CAQECVHGRCVAPNQCCAPGWRGGDCSSECAPGMWGPQCDKFCHCGNNSCDPKSGT
CFCPSGLQPPNCLQPCPAGHYGPACQFDCQCYGASCDPQDGACFCPPGRAGPSCNVPC
SQGTDGFFCPRTYPCQNGGVPQGSQGSQSCSPPGWMGVICSLPCPEGFHGPNCTQECRC
HNGGLCDRFTGQCHCAPGYIGDRCQEECPVGRFGQDCAETCDCAPGARCFPANGACLC
EHGFTGDRCTERLCPDGRYGLSCQEPCTCDPEHSLSCHPMHGECSQPGWAGLHCNES
CPQDTHGPGCQEHCLCLHGGLCLADSGLCRCAPGYTGPHCANLCPDPTYGINCSSRCS
CENAIACSPIDGTCICKEGWQRGNCSVPCPLGTWGFNCNASCQCAHDGVCSPTGACT
CTPGWHGAHCQLPCPKGQFGEGCASVCDCHSDGCDPVHGGQCRCQAGWMGTRCHLPCP
EGFWGANCSNTCTCKNGGTCVSENGNCVCAPGFRGPSCQRPCPPGRYKRCVQCKCNN

NHSSCHPSDGTCSCLAGWTGPDCEACPPGHWGLKCSQLCQCHHGGTCHPQDGSICIT
PGWTGPNCLEGCPPRMFGVNCSQLCQCDLGEMCHPETGACVCPPGHSGADCKMGSQES
FTIMPTSPVTHNSLGAVIGIAVLGTLVVALIALFIGYRQWQKGKEHEHLAVAYSTGRL
DGSDYVMPDVSPSYSHYYSNPYSYHTLSQCSPNPPPPNKVPGSQLFVSSQAPERPSRAH
GRENHVTLPADWKHRRPHERGASHLDRSYSCSYSHRNGPGPFCHKGPISSEGLGASV
MSLSSENPYATIRDLPSLPGEPRSGYVEMKGPPSVSPPRQSLHLRDRQQRQLQPQRD
SGTYEQSPSLSHNEESLGSTPPLPPGLPPGHYDSPKNSHIPGHYDLPPVRHPPSPPSR
RQDR"

Query Match 18.9%; Score 646.2; DB 11; Length 3556;
Best Local Similarity 57.1%; Pred. No. 6.8e-173;
Matches 1239; Conservative 0; Mismatches 918; Indels 12; Gaps 3;

Qy	731	GAGTGTGTGCATCACGTCACTGGAGAATGCTCTTGGCCCTTCTGGCTGGATGGGCACAGTGT	790
Db	914	GTGTTTCCTCAGGGGCTCCCAAGGCTCCTGCAGCTGCCACCAGGGCTGGATGGGTGTCATTT	973
Qy	791	GTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCC	850
Db	974	GTTCCCTGCCATGCCAGAGGGTTTCCATGGACCCAACGTGTACTCAGGAATGTCGCTGCC	1033
Qy	851	ATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAG	910
Db	1034	ACAACGGTGGCCTCTGTGACAGGTTTACTGGGCAGTGCCACTGTGCTCCTGGCTATATCG	1093
Qy	911	GGGAACGGTGCCAGGATGAGTGTCTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCT	970
Db	1094	GGGATCGGTGCCAAGAAGAGTGCCCCGTGGGCCGCTTCGGTCAAGACTGTGCTGAGACCT	1153
Qy	971	GCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAG	1030
Db	1154	GTGACTGTGCTCCTGGCGCCCGTTGCTTTCCTGCTAATGGCGCGTGTCTGTGTGAACATG	1213
Qy	1031	GCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTACGGCATCAAAT	1090
Db	1214	GCTTCACAGGCGACCGCTGCACTGAGCGCCTCTGTCCGGATGGCCGCTATGGTCTGAGCT	1273
Qy	1091	GTGACAAACGGTGTCCCTGCCACTTGAAAAACACTCATAGCTGTACCCCATGTCTGGAG	1150
Db	1274	GCCAGGAGCCCTGCACCTGCGACCCAGAACACAGTCTCAGCTGCCACCCGATGCACGGCG	1333
Qy	1151	AGTGTGCCTGCAAGCCGGGCTGGTCAAGACTCTACTGTAATGAGACATGTTCTCCTGGAT	1210
Db	1334	AGTGTCTCCTGCCAGCCAGGTTGGGCGGGCCTCCACTGCAACGAGAGCTGCCCTCAGGACA	1393
Qy	1211	TCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTG	1270
Db	1394	CGCATGGCCCGGGCTGCCAGGAGCACTGCCTCTGTCTGCACGGAGGGCTCTGCCTTGCCG	1453
Qy	1271	TGACTGGAAAGTGACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCC	1330
Db	1454	ACAGCGGCCTCTGCCGGTGCAGCGCCGGGATACACGGGACCTCACTGCGCTAACCTATGTC	1513
Qy	1331	CTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCT	1390
Db	1514	CACCGGACACTTACGGGATCAACTGTTCCCTCCGCTGCTCCTGTGAAAATGCCATTGCCT	1573
Qy	1391	GCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCA	1450
Db	1574	GCTCTCCCATCGACGGCACGTGCATCTGCAAGGAAGGTTGGCAGCGTGGTAACTGCTCTG	1633
Qy	1451	TCAGATGTCCCACTGGCACATGGGGCTTTGGCTGTAACCTTAACATGCCAGTGCCTCAACG	1510
Db	1634	TTCCCTGTCCCTTGGCACCTGGGGCTTCAATTGCAATGCCAGTTGCCAGTGTGCCACG	1693
Qy	1511	GGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGA	1570
Db	1694	ACGGAGTCTGCAGCCCCAACTGGAGCCTGTACTTGCACCCCTGGGTGGCATGGTGCTC	1753

Qy	1571	AATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACGTGCTGAGCGCTGCGACT	1630
Db	1754	ACTGCCAGCTTCCCTGCCCCAAGGGACAGTTTGCTGAAGGCTGTGCCAGTGTCTGTGACT	1813
Qy	1631	GCAGCCACGCAGATGGCTGCCACCCTACCACGGGGCATTGCCGCTGCCTCCCGGGATGGT	1690
Db	1814	GTGACCACTCTGATGGCTGTGACCCTGTTCATGGACAGTGCCGATGTCAGGCTGGTTGGA	1873
Qy	1691	CAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACCTGCTCCCTGC	1750
Db	1874	TGGGCACACGCTGCCACCTGCCTTGCCCGAGGGCTTTTGAGGAGCCAACCTGCAGTAACA	1933
Qy	1751	CCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCAC	1810
Db	1934	CCTGTACCTGCAAGAATGGTGGTACCTGTGTGTCTGAGAATGGCAACTGCGTGTGCGCAC	1993
Qy	1811	CAGGCTTCCGAGGCACCACTTGTCTGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCT	1870
Db	1994	CAGGGTTCCGAGGCCCCCTCCTGCCAGAGGCCCTGCCCGCTGGTCGCTATGGCAAACGCT	2053
Qy	1871	GCAGCCAGACATGCCCACAGTGCCTTACAGCAGCGGGCCCTGCCACCACATCACCGGCC	1930
Db	2054	GTGTGCA-----ATGCAAGTGTAAACAACCAATTCTTCCCTGCCACCCATCGGACGGGA	2107
Qy	1931	TGTGTGACTGCTTGCCTGGCTTACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCA	1990
Db	2108	CCTGCTCCTGCCTGGCGGGCTGGACAGGCCCTGACTGCTCCGAGGCATGTCCCCCAGGCC	2167
Qy	1991	GATTTGGGAAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCA	2050
Db	2168	ACTGGGGACTCAAATGCTCCCAACTCTGCCAGTGTGCATCATGGTGGGACCTGCCACCCCC	2227
Qy	2051	TTGACAGATCTTGTCACTGTACCCCGTTGGATTGGCAGTGAAGTGTCTCAACCATGTC	2110
Db	2228	AGGATGGGAGCTGTATCTGCACGCCAGGCTGGACTGGACCCAACCTGCTTGAAGGCTGCC	2287
Qy	2111	CACCTGCCCCTGGGGCCCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCT	2170
Db	2288	CACCAAGAATGTTTGGTGTCAACTGCTCCAGCTATGTGAGTGTGATCTCGGAGAGATGT	2347
Qy	2171	GCAGCGCTACGATGGGGAATGTAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTC	2230
Db	2348	GCCACCCAGAGACTGGGGCTTGTGTCTGTCCCCCAGGACACAGTGGTGCAGACTGCAAAA	2407
Qy	2231	AGAGATGTC	2239
Db	2408	TGGGAAGCC	2416

RESULT 8

LOCUS CD803668 789 bp mRNA linear EST 15-JUL-2003

DEFINITION UI-M-GV0-chv-b-08-0-UI.r1 NIH_BMAP_GV0 Mus musculus cDNA clone
IMAGE:30544831 5', mRNA sequence.

ACCESSION CD803668

VERSION CD803668.1 GI:32462494

KEYWORDS EST.

Db	61	CAACGTATGCAGCCACTGGGGAAAGCTACTCGGTGACTGTGCAGGAGTCGTATCCACATC	120
Qy	152	CCTTTGATCAAATTTACTACACGAGCTGCACTGACATTCTAAACTGGTTTAAATGCACGC	211
Db	121	CCTTCGATCAGATCTACTACACAAGCTGCACCGACATCCTGAACTGGTTTAAATGCACAC	180
Qy	212	GGCACAGAGTCAGCTATCGGACAGCCTATCGACATGGG-GAGAAGACTATGTATAGGCGC	270
Db	181	GGCACAGAATCAGCTACCGGACAGCCTACCGCCACGGGNGAGAAAACCATGTATAGACGC	240
Qy	271	AAGTCTCAGTGTTGTCCTGGATTTTATGAAAGCGGGGAAATGTGTGTCCCCACTGTGCT	330
Db	241	AAATCCCAGTGTTGCCAGGATTTTATGAAAGCCGAGACATGTGTGTCCCTCACTGTGCT	300
Qy	331	GATAAATGTGTCCATGGTCGCTGTATTGCTCCAAACACCTGTCAGTGTGAGCCTGGCTGG	390
Db	301	GATAAATGTGTCCATGGTCGCTGCATTGCTCCAAACACCTGTCAGTGTGAGCCTGGCTGG	360
Qy	391	GGAGGGACCAACTGCTCCAGTGCCTGCGATGGTGATCACTGGGGTCCCCACTGCACCAGC	450
Db	361	GGTGGGACCAACTGTAGCAGTGCTTGTGATGGTGATCACTGGGGCCTCACTGCAGCAGC	420
Qy	451	CGGTGCCAGTGCAAAAATGGGGCTCTGTGCAACCCCATCACCGGGCTTGCCACTGTGCT	510
Db	421	CGATGCCAGTGCAAAAACAGAGCTTTGTGTAACCCCATCACCGGTGCTTGCCACTGCGCT	480
Qy	511	GCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGCAGGGCACCTATGGTAACGAC	570
Db	481	GCGGGCTACCGGGGATGGCGCTGCGAGGACCGTTGTGAACAGGGCACGTACGGTAACGAC	540
Qy	571	TGTCATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCGACCACGTACGCGGGGAATGC	630
Db	541	TGTCACCAAAGATGCCAGTGTCAGAATGGAGCGACCTGTGACCACATCACTGGGGGAATGC	600
Qy	631	CGCTGCCCACCAGGATACACCGGAGCCTTCTGTGAGGATCTTTGTCCTCCTGGTAAACAT	690
Db	601	CGTTGTTCACTGGGTACACTGGAGCCTTCTGTGAGGATCTTTGTCCTCCTGGCANACAT	660
Qy	691	GGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAATGGAGGAGTGTGTCATCACGTCACT	750
Db	661	GGTCCACATTGTGAGCAGAGGTGTCCCTGCCANAATGGGGGTGTGTGCCACCATGTCACT	720
Qy	751	GGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGTGTGGTCAGCCTTGCCCCGAG	810
Db	721	GGAGAGTGCTCTTGCCCTTCTGGNTGGATGGGCACAGTGTGTGGTCAGCCCTGCCCTGAN	780
Qy	811	GGTCGCTTT	819
Db	781	GGTCGCTTT	789

RESULT 9

BU215784

LOCUS BU215784 921 bp mRNA linear EST 25-NOV-2002

DEFINITION 603106167F1 CSEQCHN04 Gallus gallus cDNA clone ChEST45f18 5', mRNA sequence.

ACCESSION BU215784

VERSION BU215784.1 GI:25394329
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 921)
 AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
 TITLE A Comprehensive Collection of Chicken cDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 22335534
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 FEATURES Location/Qualifiers
 source 1. .921
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="ChEST45f18"
 /tissue_type="whole embryo"
 /dev_stage="20-21"
 /lab_host="DH10B"
 /clone_lib="CSEQCHN04"
 /note="Organ: whole embryo; Vector: pBluescript II KS(+);
 Site_1: EcoRI; Site_2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."
 ORIGIN
 Query Match 15.7%; Score 539; DB 13; Length 921;
 Best Local Similarity 75.6%; Pred. No. 1.7e-142;
 Matches 695; Conservative 0; Mismatches 220; Indels 4; Gaps 2;
 Qy 787 GTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAG 846
 || ||||| || || ||||| || || ||||| || || |||||
 Db 3 GTCTGTGGTCAGCCTTGTCCTGAAGGTCGTTATGGGAATAACTGTTCCCAGGAGTGTCTAG 62

Qy	847	TGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTTCATTGCAGTCCAGGATAC	906
Db	63		122
Qy	907	ACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAG	966
Db	123		182
Qy	967	ACCTGCCAGTGTGTCAACGGAGGGGAAGTGTTACCACGTGAGCGGCGCATGCCTCTGTGAA	1026
Db	183		242
Qy	1027	GCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTACGGCATC	1086
Db	243		302
Qy	1087	AAATGTGACAAACGGTGTCCCTGCCACTTGGAAAACTCATAGCTGTACCCCATGTCT	1146
Db	303		362
Qy	1147	GGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCT	1206
Db	363		422
Qy	1207	GGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGAC	1266
Db	423		482
Qy	1267	AGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCA	1326
Db	483		542
Qy	1327	TGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCA	1386
Db	543		602
Qy	1387	GTCTGCTCTCCTGTGGACGGGTCTTGTAATTGCAAGGCAGGCTGGCACGGGGTGGACTGC	1446
Db	603		662
Qy	1447	TCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACTTAACATGCCAGTGCCTC	1506
Db	663		722
Qy	1507	AACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGG	1566
Db	723		782
Qy	1567	GAGAAATGCGAACTTCCCT--GCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCT	1624
Db	783		842
Qy	1625	GCG--ACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCC	1682
Db	843		902
Qy	1683	GGGATGGTCAGGTGTCCAC	1701

Db | | | | | | |
903 AGATGGTCCGGTTTCCCTC 921

RESULT 10
CD802967

LOCUS CD802967 755 bp mRNA linear EST 15-JUL-2003

DEFINITION UI-M-GV0-ch1-j-22-0-UI.r1 NIH_BMAP_GV0 Mus musculus cDNA clone
IMAGE:30543885 5', mRNA sequence.

ACCESSION CD802967

VERSION CD802967.1 GI:32461793

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 755)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>

This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES

source

Location/Qualifiers

1. .755

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:30543885"

/tissue_type="whole brain"

/dev_stage="1,5, and 15 days newborn"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_GV0"

/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;

Site_2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated with

EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is CGAAGTGAAT. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

ORIGIN

Query Match 15.5%; Score 530.4; DB 14; Length 755;
 Best Local Similarity 84.5%; Pred. No. 4.4e-140;
 Matches 594; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

```

Qy      1 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 60
      |||  |||||  |  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      52 ATGGCGATTTCTTCAAGTTCGTGCCTGGGCCTCATCTGCTCACTGCTCTGTCACTGGGTG 111

Qy      61 GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC 120
      |||||  |||||  |  |||||  ||  |||||  ||  ||  ||  |||||  |||||  |||||
Db      112 GGGACAGCATCCTCCCTGAACCTGGAAGACCCCAACGTATGCAGCCACTGGGAAAGCTAC 171

Qy      121 TCAGTGA CTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC 180
      ||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      172 TCGGTGA CTGTGCAGGAGTCGTATCCACATCCCTTCGATCAGATCTACTACACAAGCTGC 231

Qy      181 ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT 240
      ||  |||||  ||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      232 ACCGACATCCTGAAC TGGTTTAAATGCACACGCGGCACAGAATCAGCTACCGGACAGCCTAC 291

Qy      241 CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA 300
      ||  ||  |||||  ||  |||||  |||||  ||  |||||  ||  |||||  |||||
Db      292 CGCCACGGGGAGAAAACCATGTATAGACGCAAATCCCAGTGTTGCCCAGGATTTTATGAA 351

Qy      301 AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT 360
      |||  |  ||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      352 AGCCGAGACATGTGTGTCCCTCACTGTGCTGATAAATGTGTCCATGGTCGCTGCATTGCT 411

Qy      361 CCAAACACCTGTCACTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT 420
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      412 CCAAACACCTGTCACTGTGAGCCTGGCTGGGGTGGGACCAACTGTAGCAGTGCTTGTGAT 471

Qy      421 GGTGATCACTGGGGTCCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC 480
      |||||  |||||  ||  |||||  |||||  |||||  |||||  |||||  |||||
Db      472 GGTGATCACTGGGGCCTCACTGCAGCAGCCGATGCCAGTGCAAAAACAGAGCTTTGTGT 531

Qy      481 AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC 540
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      532 AACCCCATCACCGGTGCTTGCCACTGCGCTGCGGGCTACCGGGGATGGCGCTGCGAGGAC 591

Qy      541 CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA 600
      ||  |||||  |||||  ||  |||||  |||||  ||  |||||  |||||  |||||
Db      592 CGTTGTGAACAGGGCACGTACGGTAACGACTGTCACCANAGATGCCAGTGTCAGAATGGA 651

Qy      601 GCCACCTGCGACCACGTACGCGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC 660
      ||  |||||  |||||  |||||  |||||  |||||  ||  |||||  |||||  |||||
Db      652 GCGACCTGTGACCACATCACTGGGGAATGCCGTTGTTACCTGNGTACACTNGGAGCTTC 711

Qy      661 TGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTG 703
      |||||  |||||  |||||  ||  |  ||  |||  |
Db      712 TGTGAGGATCTTTGTCTCCTGGCAACATGGTCACATTGTGAG 754
  
```

LOCUS BU056532 643 bp mRNA linear EST 26-AUG-2002
 DEFINITION UI-M-FO0-cab-n-24-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
 IMAGE:6409175 5', mRNA sequence.
 ACCESSION BU056532
 VERSION BU056532.1 GI:22496609
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 643)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.
 FEATURES Location/Qualifiers
 source 1. .643
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6409175"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_F00"
 /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is TGAGAGAGCC. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."
 ORIGIN
 Query Match 14.9%; Score 508.6; DB 13; Length 643;
 Best Local Similarity 86.9%; Pred. No. 7.2e-134;
 Matches 559; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 2944)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers
source 1. .2944
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C230074H09"
/db_xref="MGI:2415973"
/db_xref="taxon:10090"
/clone="C230074H09"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
CDS 154..1053
/note="unnamed protein product; MEGF11 PROTEIN (KIAA1781)
homolog [Homo sapiens] (SPTR|Q96KG6, evidence: FASTY,
85.9%ID, 85.4%length, match=2363)
putative"
/codon_start=1
/protein_id="BAC33471.1"
/db_xref="GI:26339600"
/translation="MAPSAVGLLVFLQAAALALNPEDPNVCSHWESYAVTVQESYAHF
FDQIYYTRCADILNWFKCTRHRISYKTAYRRGLRTMYRRRSQCCPGYYENGDFCIPLC
TEECMHGRCVSPDTCHEPGWGGPDCSSGCDSEHWGPHCSNRCQCQNGALCNPITGAC
VCAPGFRGWRCEELCAPGTHGKGCQLLCQCHHGASCDPRTGECCLCAPGYTGVYCEELC
PPGSHGAHCELRCPQNGGTCHHITGECACPPGWTGAVCAQPCPPGTFGQNCSDCPC
HHGGQCDHVTGQCHCTAGYMGDR"

ORIGIN

Query Match 13.7%; Score 470; DB 11; Length 2944;
Best Local Similarity 72.2%; Pred. No. 2.3e-122;
Matches 611; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

```
Qy      74 CTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACTCAGTGACTGTGC 133
      ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db      206 CTCTGAACCCTGAAGACCCCAATGTGTGTAGCCACTGGGAGAGCTATGCCGTGACTGTGC 265

Qy      134 AAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGCACTGACATTCTAA 193
      | |||| | | ||| ||||| | ||||| | ||||| | ||| | ||||| | |
Db      266 AGGAGTCTTATGCACACCCCTTTGATCAGATCTACTACACACGATGTGCAGACATCCTCA 325

Qy      194 ACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTATCGACATGGGGAGA 253
      ||||| | | || | ||||| | ||||| | ||| | || | | ||
Db      326 ACTGGTTCAAGTGTACCGGCACCGGATCAGCTATAAGACCGCGTATAGGCGCGGCCTCC 385

Qy      254 AGACTATGTATAGGCGCAAGTCTCAGTGTGTCTGGATTTTATGAAAGCGGGGAAATGT 313
```

Db	386	GGACCATGTACCGGCGGAGGTCCCAATGCTGCCCTGGCTACTATGAGAACGGAGACTTCT	445
Qy	314	GTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCTCCAAACACCTGTC	373
Db	446	GCATTCTCTGTGTACCGAGGAGTGCATGCACGGCCGCTGTGTCTCTCCCGATACCTGCC	505
Qy	374	AGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGATGGTGATCACTGGG	433
Db	506	ACTGTGAGCCTGGATGGGGAGGCCCTGACTGCTCCAGCGGCTGTGACAGCGAGCACTGGG	565
Qy	434	GTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGCAACCCCATCACCG	493
Db	566	GTCCCCACTGCAGCAACCGGTGTCAGTGTGACAACGGCGCCCTGTGCAACCCATCACCG	625
Qy	494	GGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGCAGG	553
Db	626	GCGCCTGCGTGTGCGCCCCGGGCTTCCGAGGCTGGCGCTGTGAGGAACCTGCGCTCCTG	685
Qy	554	GCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCGACC	613
Db	686	GTA CTCACGGCAAGGGCTGCCAGCTGCTCTGTCA GTGCCACCATGGCGCCAGCTGTGACC	745
Qy	614	ACGTCACGGGGGAATGCCGCTGCCCACCAGGATACACCGAGCCTTCTGTGAGGATCTTT	673
Db	746	CGCGCACTGGCGAGTGCCTCTGCGCTCCTGGCTACACAGGCGTTTACTGTGAGGAGCTGT	805
Qy	674	GTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAAATGGAGGAG	733
Db	806	GCCCCCTGGGAGCCATGGAGCTCACTGTGAGCTGCGCTGCCCCTGCCAGAATGGAGGCA	865
Qy	734	TGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGTGTG	793
Db	866	CCTGCCACCACATCACTGGCGAATGTGCCTGCCCTCCAGGCTGGACGGGAGCAGTGTGTG	925
Qy	794	GTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCCATA	853
Db	926	CCCAGCCCTGCCCTCCAGGACCTTTGGCCAGA ACTGTAGCCAGGACTGTCCCTGCCACC	985
Qy	854	ATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAGGGG	913
Db	986	ATGGAGGCCAGTGTGACCATGTGACTGGACAATGCCACTGTACAGCTGGATACATGGGGG	1045
Qy	914	AACGGT	919
Db	1046	ACAGGT	1051

RESULT 13

BG828819

LOCUS BG828819 755 bp mRNA linear EST 22-MAY-2001
 DEFINITION 602751395F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4904255 5',
 mRNA sequence.

ACCESSION BG828819

VERSION BG828819.1 GI:14176418

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 755)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1803 row: m column: 24
High quality sequence stop: 707.

FEATURES
source Location/Qualifiers
1. .755
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4904255"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_17"
/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 13.3%; Score 456.6; DB 12; Length 755;
Best Local Similarity 98.8%; Pred. No. 6.7e-119;
Matches 481; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy 1 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 60
|||||

Db 268 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 327

Qy 61 GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC 120
|||||

Db 328 GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC 387

Qy 121 TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC 180
|||||

Db 388 TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC 447

Qy 181 ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT 240
|||||

Db 448 ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT 507

Qy 241 CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA 300
 |||
 Db 508 CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA 567
 Qy 301 AGCGGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTC-GCTGTATTGC 359
 |||
 Db 568 AGCGGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTCGGCTGTATTGC 627
 Qy 360 TCCAAACACCTGTCAGTGTCAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCCA 419
 |||
 Db 628 TCCAAACACCTGTCAGTGTCAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCCA 687
 Qy 420 TGGTGATCACTGGGGTCCCCACTGCACCAGCCGGTG-CCAGTGCAAAAATGGGGCTCTGT 478
 |||
 Db 688 TGGTGATCACTGGAGTCCCAATGCACCAGCCGGTGCCCAGTGCAAACATGGGGCTCTGT 747
 Qy 479 GCAACCC 485
 ||
 Db 748 GCCACCC 754

RESULT 14

BM719978

LOCUS BM719978 565 bp mRNA linear EST 01-MAR-2002

DEFINITION UI-E-EJ0-ahu-i-16-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone
 UI-E-EJ0-ahu-i-16-0-UI 5', mRNA sequence.

ACCESSION BM719978

VERSION BM719978.1 GI:19038910

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 565)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa

DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

FEATURES Location/Qualifiers

source 1. .565

/organism="Homo sapiens"


```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-ahu-i-16-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-EJ0 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,
AATGCCGCAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."
```

ORIGIN

```

Query Match          12.6%;  Score 433;  DB 12;  Length 565;
Best Local Similarity 100.0%;  Pred. No. 3.3e-112;
Matches 433;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```

Qy      2800 ACCCAGTGTGCCACATCCCCTCACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCA 2859
          |||
Db        1 ACCCAGTGTGCCACATCCCCTCACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCA 60

Qy      2860 AAAACAATCAACTGTTTGTGAATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTG 2919
          |||
Db        61 AAAACAATCAACTGTTTGTGAATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTG 120

Qy      2920 GGGGACTGCACTGGGACATTGCCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTC 2979
          |||
Db       121 GGGGACTGCACTGGGACATTGCCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTC 180

Qy      2980 GGTGCTTTTGGACTTGACAGAAGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAAT 3039
          |||
Db       181 GGTGCTTTTGGACTTGACAGAAGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAAT 240

Qy      3040 TCTGAATATAATTCAAGTAACTGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATT 3099
          |||
Db       241 TCTGAATATAATTCAAGTAACTGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATT 300

Qy      3100 AAAGACCCACCTGTACTTATCCCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCG 3159
          |||
Db       301 AAAGACCCACCTGTACTTATCCCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCG 360

Qy      3160 CCGGCACGAAGAGATTCCCATATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAAT 3219
```

```

          |||
Db      361 CCGGCACGAAGAGATTCCCCATATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAAT 420
          |||
Qy      3220 GTCTATGAAGTTG 3232
          |||
Db      421 GTCTATGAAGTTG 433

```

RESULT 15

BM676825/c

LOCUS BM676825 598 bp mRNA linear EST 27-FEB-2002

DEFINITION UI-E-EJ0-ahu-i-16-0-UI.s2 UI-E-EJ0 Homo sapiens cDNA clone
UI-E-EJ0-ahu-i-16-0-UI 3', mRNA sequence.

ACCESSION BM676825

VERSION BM676825.1 GI:18986721

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 598)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa

DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa

Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).

Seq primer: M13 Forward

POLYA=Yes.

FEATURES Location/Qualifiers

source 1. .598
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-ahu-i-16-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-EJ0 is a subtracted cDNA library constructed

according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGCAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG_TISSUE=human fetal eyes
TAG_LIB=UI-E-EJ0
TAG_SEQ=AGAATCAAGA"

ORIGIN

Query Match 12.6%; Score 432; DB 12; Length 598;
Best Local Similarity 98.9%; Pred. No. 6.6e-112;
Matches 435; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Qy      2793 CACGCTCACCCAGTGTGCCACATCCCCTCACGTCAACAACAGGGACAGGATGACTGTCAC 2852
          ||||  ||||||||||||||||||||||||||||||||||||||||||||
Db      596 CACGAGGACCCAGTGTGCCACATCCCCTCACGTCAACAACAGGGACAGGATGACTGTCAC 537

Qy      2853 GAAGTCAAAAAACAATCAACTGTTTGTGAATCTTAAAAATGTGAACCCTGGGAAGAGAGG 2912
          ||||||||||||||||||||||||||||||||||||||||||||
Db      536 GAAGTCAAAAAACAATCAACTGTTTGTGAATCTTAAAAATGTGAACCCTGGGAAGAGAGG 477

Qy      2913 CCCTGTGGGGGACTGCACTGGGACATTGCCGGCTGACTGGAAACATGGCGGCTACCTCAA 2972
          ||||||||||||||||||||||||||||||||||||||||||||
Db      476 CCCTGTGGGGGACTGCACTGGGACATTGCCGGCTGACTGGAAACATGGCGGCTACCTCAA 417

Qy      2973 CGAGCTCGGTGCTTTTGGACTTGACAGAAGCTATATGGGAAAATCCTTAAAGACCTGGG 3032
          ||||||||||||||||||||||||||||||||||||||||||||
Db      416 CGAGCTCGGTGCTTTTGGACTTGACAGAAGCTATATGGGAAAATCCTTAAAGACCTGGG 357

Qy      3033 AAAGAATTCTGAATATAATTCAAGTAAGTCTCCCTAAGCAGTTCTGAGAACCCATATGC 3092
          ||||||||||||||||||||||||||||||||||||||||||||
Db      356 AAAGAATTCTGAATATAATTCAAGTAAGTCTCCCTAAGCAGTTTCTGAGAACCCATATGC 297

Qy      3093 CACTATTAAAGACCCACCTGTACTTATCCCGAAAAGCTCAGAGTGTGGTTATGTGGAGAT 3152
          ||||||||||||||||||||||||||||||||||||||||||||
Db      296 CACTATTAAAGACCCACCTGTACTTATCCCGAAAAGCTCAGAGTGTGGTTATGTGGAGAT 237

Qy      3153 GAAATCGCCGGCACGAAGAGATTCCCCATATGCAGAGATCAATAACTCAACTTCAGCCAA 3212
          ||||||||||||||||||||||||||||||||||||||||||||
Db      236 GAAATCGCCGGCACGAAGAGATTCCCCATATGCAGAGATCAATAACTCAACTTCAGCCAA 177

Qy      3213 CAGGAATGTCTATGAAGTTG 3232
          ||||||||  ||||||||
Db      176 CAGGAATGTTTATGAAGTTG 157

```

Search completed: March 30, 2004, 08:33:24
Job time : 8353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 01:07:41 ; Search time 13015 Seconds
(without alignments)
11399.401 Million cell updates/sec

Title: US-10-092-390-1
Perfect score: 3423
Sequence: 1 atgggttattttctttgaactc.....gcagcagcagcagcagtgaatga 3423

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*

28: em_un:*
 29: em_vi:*
 30: em_htg_hum:*
 31: em_htg_inv:*
 32: em_htg_other:*
 33: em_htg_mus:*
 34: em_htg_pln:*
 35: em_htg_rod:*
 36: em_htg_mam:*
 37: em_htg_vrt:*
 38: em_sy:*
 39: em_htgo_hum:*
 40: em_htgo_mus:*
 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	% Query						Description
	No.	Score	Match	Length	DB	ID	
	1	3421.4	100.0	7522	6	BD185216	BD185216 Novel gen
	2	3421.4	100.0	7522	9	AB058676	AB058676 Homo sapi
	3	2274.8	66.5	2952	9	AK123568	AK123568 Homo sapi
	4	1692.4	49.4	2267	9	BC020198	BC020198 Homo sapi
	5	1425.8	41.7	1448	6	AR217554	AR217554 Sequence
	6	1425.8	41.7	1448	6	BD080296	BD080296 Tango-71,
	7	1200.6	35.1	5702	9	AB058677	AB058677 Homo sapi
	8	1040.8	30.4	5278	10	AK122555	AK122555 Mus muscu
	9	742	21.7	3281	9	HSM805375	AL834326 Homo sapi
	10	649.4	19.0	4290	10	AF444274	AF444274 Mus muscu
	11	648.6	18.9	4539	10	AF461685	AF461685 Mus muscu
	12	646.2	18.9	4482	10	AF440279	AF440279 Mus muscu
	13	600.6	17.5	3574	6	AX492979	AX492979 Sequence
	14	548.2	16.0	660	10	RATORFD	L41686 Rattus norv
	15	534.8	15.6	632	6	AX079681	AX079681 Sequence
	16	515.4	15.1	3503	10	BC042490	BC042490 Mus muscu
	17	456.8	13.3	4470	9	AK074121	AK074121 Homo sapi
	18	400.8	11.7	5523	10	AB011532	AB011532 Rattus no
	19	396.8	11.6	7334	6	AX817303	AX817303 Sequence
	20	393	11.5	4835	6	AX817305	AX817305 Sequence
	21	387	11.3	5178	6	AX704753	AX704753 Sequence
	22	386.8	11.3	4733	6	AX817299	AX817299 Sequence
	23	383.8	11.2	7319	6	AX817297	AX817297 Sequence
	24	383.8	11.2	7337	6	AX817295	AX817295 Sequence
	25	365.4	10.7	3391	10	BC031402	BC031402 Mus muscu
	26	365.4	10.7	3402	10	BC039980	BC039980 Mus muscu
	27	362.8	10.6	2629	10	BC058571	BC058571 Mus muscu
	28	354.8	10.4	4501	9	AB011539	AB011539 Homo sapi
c	29	250.8	7.3	146206	9	AC026800	AC026800 Homo sapi
	30	250.8	7.3	152765	2	AC008566	AC008566 Homo sapi
c	31	250.8	7.3	175144	2	AC010415	AC010415 Homo sapi
c	32	250.8	7.3	217221	9	AC008682	AC008682 Homo sapi
c	33	243	7.1	48202	2	AC012911	AC012911 Drosophil

	34	243	7.1	172175	3	AC010038	AC010038	Drosophil
c	35	243	7.1	177583	3	AC105264	AC105264	Drosophil
c	36	243	7.1	257867	3	AC005557	AC005557	Drosophil
c	37	243	7.1	303191	3	AE003472	AE003472	Drosophil
	38	241.8	7.1	192282	9	AC010424	AC010424	Homo sapi
	39	239.8	7.0	2054	9	AK098809	AK098809	Homo sapi
	40	216	6.3	3764	3	AF332568	AF332568	Caenorhab
	41	197.2	5.8	185768	10	AC102173	AC102173	Mus muscu
c	42	192.4	5.6	194593	2	AC139989	AC139989	Rattus no
	43	192.4	5.6	226956	10	AC095558	AC095558	Rattus no
	44	192.4	5.6	231950	2	AC094759	AC094759	Rattus no
c	45	180	5.3	209476	2	AC122117	AC122117	Mus muscu

ALIGNMENTS

RESULT 1

BD185216

LOCUS BD185216 7522 bp DNA linear PAT 17-JUN-2003

DEFINITION Novel genes and proteins encoded by the genes.

ACCESSION BD185216

VERSION BD185216.1 GI:31877416

KEYWORDS JP 2002345493-A/59.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 7522)

AUTHORS Ohara,O., Nagase,T. and Nakajima,D.

TITLE Novel genes and proteins encoded by the genes

JOURNAL Patent: JP 2002345493-A 59 03-DEC-2002;

KAZUSA DNA RESEARCH INSTITUTE

COMMENT OS Homo sapiens (human)

PN JP 2002345493-A/59

PD 03-DEC-2002

PF 26-FEB-2002 JP 2002049046

PI OSAMU OHARA,TAKAHIRO NAGASE,DAISUKE NAKAJIMA

PC C12N15/09,C07K14/47,C07K14/54,C12N15/00

CC Novel genes and proteins encoded by the genes FH Key

Location/Qualifiers

FT CDS (48)..(3623).

FEATURES Location/Qualifiers

source 1..7522

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 3421.4; DB 6; Length 7522;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 60

|||||

Db 204 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 263

Qy	61	GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC	120
Db	264	GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC	323
Qy	121	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	180
Db	324	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	383
Qy	181	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	240
Db	384	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	443
Qy	241	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	300
Db	444	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	503
Qy	301	AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	360
Db	504	AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	563
Qy	361	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	420
Db	564	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	623
Qy	421	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	480
Db	624	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	683
Qy	481	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	540
Db	684	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	743
Qy	541	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	600
Db	744	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	803
Qy	601	GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	660
Db	804	GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	863
Qy	661	TGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Db	864	TGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	923
Qy	721	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Db	924	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	983
Qy	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Db	984	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	1043
Qy	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Db	1044	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	1103
Qy	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960

Db	1104	 GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	1163
Qy	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1020
Db	1164	 GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1223
Qy	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTAC	1080
Db	1224	 TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTAC	1283
Qy	1081	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGAAAACTCATAGCTGTCACCCC	1140
Db	1284	 GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGAAAACTCATAGCTGTCACCCC	1343
Qy	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Db	1344	 ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1403
Qy	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Db	1404	 TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1463
Qy	1261	TGTGACAGTGTGACTTGAAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Db	1464	 TGTGACAGTGTGACTTGAAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1523
Qy	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAAT	1380
Db	1524	 ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAAT	1583
Qy	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTAAGTGAAGGCAGGCTGGCAGGGGTG	1440
Db	1584	 GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTAAGTGAAGGCAGGCTGGCAGGGGTG	1643
Qy	1441	GACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1500
Db	1644	 GACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1703
Qy	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1560
Db	1704	 TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1763
Qy	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1620
Db	1764	 CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1823
Qy	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1680
Db	1824	 CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1883
Qy	1681	CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1740
Db	1884	 CCCGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1943
Qy	1741	TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCTGATGATGGCATCTGC	1800

Db	1944	TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGC	2003
Qy	1801	GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTCAGAGGATCTGCTCCCCTGGTTTTTAT 	1860
Db	2004	GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTCAGAGGATCTGCTCCCCTGGTTTTTAT	2063
Qy	1861	GGGCATCGCTGCAGCCAGACATGCCCACAGTGC GTTCACAGCAGCGGGCCCTGCCACCAC 	1920
Db	2064	GGGCATCGCTGCAGCCAGACATGCCCACAGTGC GTTCACAGCAGCGGGCCCTGCCACCAC	2123
Qy	1921	ATCACCGGCCTGTGTGACTGCTTG CCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT 	1980
Db	2124	ATCACCGGCCTGTGTGACTGCTTG CCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT	2183
Qy	1981	CCCAGTGGCAGATTTGGGAAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC 	2040
Db	2184	CCCAGTGGCAGATTTGGGAAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC	2243
Qy	2041	TGTAACCCCATTTGACAGATCTTGT CAGTGTTACCCCGGTTGGATTGGCAGTGACTGCTCT 	2100
Db	2244	TGTAACCCCATTTGACAGATCTTGT CAGTGTTACCCCGGTTGGATTGGCAGTGACTGCTCT	2303
Qy	2101	CAACCATGTCCACCTGCCCCACTGGGGCCCAA ACTGCATCCACACGTGCAACTGCCATAAT 	2160
Db	2304	CAACCATGTCCACCTGCCCCACTGGGGCCCAA ACTGCATCCACACGTGCAACTGCCATAAT	2363
Qy	2161	GGAGCTTCTGCAGCGCCTACGATGGGGA ATGTAAATGCACTCCTGGCTGGACAGGGCTC 	2220
Db	2364	GGAGCTTCTGCAGCGCCTACGATGGGGA ATGTAAATGCACTCCTGGCTGGACAGGGCTC	2423
Qy	2221	TACTGCACTCAGAGATGTCCTCTAG GGTTTTATGGAAAAGATTGTGCACTGATATGCCAA 	2280
Db	2424	TACTGCACTCAGAGATGTCCTCTAG GGTTTTATGGAAAAGATTGTGCACTGATATGCCAA	2483
Qy	2281	TGTCAAAACGGAGCT'GACTGCGACC ACATTTCTGGGCAGTGTACTTGCCGCACTGGATTC 	2340
Db	2484	TGTCAAAACGGAGCT'GACTGCGACC ACATTTCTGGGCAGTGTACTTGCCGCACTGGATTC	2543
Qy	2341	ATGGGACGGCACTGTGAGCAGA AGTGCCCTTCAGGAACATATGGCTATGGCTGTGCCCAG 	2400
Db	2544	ATGGGACGGCACTGTGAGCAGA AGTGCCCTTCAGGAACATATGGCTATGGCTGTGCCCAG	2603
Qy	2401	ATATGTGATTGTCTGAACA AACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC 	2460
Db	2604	ATATGTGATTGTCTGAACA AACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC	2663
Qy	2461	CCCGGATGGAAGGGAGCGAGATGTG ATCAAGCTGGTGTTATCATAGTTGGAATCTGAAC 	2520
Db	2664	CCCGGATGGAAGGGAGCGAGATGTG ATCAAGCTGGTGTTATCATAGTTGGAATCTGAAC	2723
Qy	2521	AGCTTAAGCCGAACCACTACTGCT CTCCCTGCTGATTCCTACCAGATCGGGGCCATTGCA 	2580
Db	2724	AGCTTAAGCCGAACCACTACTGCT CTCCCTGCTGATTCCTACCAGATCGGGGCCATTGCA	2783
Qy	2581	GGCATCATCATTTCTTGTCTTAG TTGTTCTTCTTCTTCTACTGGCATTGTTTCATTATTTATAGA 	2640
Db	2784	GGCATCATCATTTCTTGTCTTAG TTGTTCTTCTTCTTCTACTGGCATTGTTTCATTATTTATAGA	2843

Qy	2641	CACAAGCAGAAGGGAAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2700
Db	2844	CACAAGCAGAAGGGAAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2903
Qy	2701	GTCGTCAATGCAGATTATACCATTTAGGAACCCTTCCTCACAGCAATGGTGGAAACGCT	2760
Db	2904	GTCGTCAATGCAGATTATACCATTTAGGAACCCTTCCTCACAGCAATGGTGGAAACGCT	2963
Qy	2761	AATAGCCACTACTTCACCAATCCCAGTTACCACACGCTCACCCAGTGTGCCACATCCCCCT	2820
Db	2964	AATAGCCACTACTTCACCAATCCCAGTTACCACACGCTCACCCAGTGTGCCACATCCCCCT	3023
Qy	2821	CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	2880
Db	3024	CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	3083
Qy	2881	AATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	2940
Db	3084	AATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	3143
Qy	2941	CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3000
Db	3144	CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3203
Qy	3001	AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAC	3060
Db	3204	AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAC	3263
Qy	3061	TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3120
Db	3264	TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3323
Qy	3121	CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCAGCAAGAGATTCCCCA	3180
Db	3324	CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCAGCAAGAGATTCCCCA	3383
Qy	3181	TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA	3240
Db	3384	TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA	3443
Qy	3241	GTGAGTGTGTCCAAGGAGTATTAGCAATAATGGGCGTCTCTCCAGGATCCATATGAC	3300
Db	3444	GTGAGTGTGTCCAAGGAGTATTAGCAATAATGGGCGTCTCTCCAGGATCCATATGAC	3503
Qy	3301	CTCCCAAAGAACAGTCACATCCCTTGTATTATGACCTGCTGCCAGTCCGAGACAGTTCA	3360
Db	3504	CTCCCAAAGAACAGTCACATCCCTTGTATTATGACCTGCTGCCAGTCCGAGACAGTTCA	3563
Qy	3361	TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGTGA	3420
Db	3564	TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGTGA	3623
Qy	3421	TGA	3423
Db	3624	TGA	3626

RESULT 2

AB058676

LOCUS AB058676 7522 bp mRNA linear PRI 10-MAY-2001

DEFINITION Homo sapiens mRNA for MEGF10 protein (KIAA1780), complete cds.

ACCESSION AB058676

VERSION AB058676.1 GI:14017776

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)

AUTHORS Nagase,T., Nakayama,M., Nakajima,D., Kikuno,R. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes. XX.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro

JOURNAL DNA Res. 8 (2), 85-95 (2001)

MEDLINE 21245130

PUBMED 11347906

REFERENCE 2 (bases 1 to 7522)

AUTHORS Nakayama,M., Nagase,T., Nakajima,D., Kikuno,R. and Ohara,O.

TITLE Direct Submission

JOURNAL Submitted (27-MAR-2001) Manabu Nakayama, Kazusa DNA Research
Institute, Department of Human Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:nmanabu@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3915,
Fax:81-438-52-3914)

FEATURES Location/Qualifiers

source

1. .7522
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pf01012"
/tissue_type="hippocampus"
/dev_stage="adult"
/note="vector:pBluescript II SK plus"

gene

1. .7522
/gene="MEGF10"

CDS

204. .3626
/gene="MEGF10"
/note="KIAA1780 protein
gene encoding protein with multiple EGF-like-domains"
/codon_start=1
/product="MEGF10 protein (KIAA1780)"
/protein_id="BAB47409.1"
/db_xref="GI:14017777"
/translation="MVISLNSCLSFI~~CL~~LLCHWIGTASPLNLEDPNVCSHWESYSVTV
QESYPHPFDQIIYTSCTDILNWFKCTRHRVSYRTAYRHGEKTMYYRRKSQCCPGFYESG
EMCVPHCADKCVHGRCIAPNTCQCEPGWGGTNCSSACGDHWGPHCTSRCQCKNGALC
NPITGACHCAAGFRGWRCEDRCEQGTYGNDCHQRCQCNQATCDHVTGECRCPPGYTG
AFCE~~DL~~CPPGKHGPQCEQRCPCQNGGVCHHVTGECSCPSGWMGTVCQGQPCPEGRFGKN
CSQECQCHNGGTCDAA~~TG~~QCHCSPGYTGERCQDECPVGTYGVLC~~AE~~TCQCVNGGKCYH
VSGAC~~L~~CEAGFAGERCEARLCPEGLYGIKDKRCPC~~HL~~ENTH~~SH~~PM~~SG~~ECACKPGWS
GLYCN~~ET~~CSPGFYGEACQ~~Q~~ICSCQNGADCD~~SV~~TGKCTCAPGFKGIDCSTPCPLGTYGI
NCSSRCGCKND~~AV~~CS~~PD~~VG~~SCT~~CKAGWHGVDCSIRCPSGTWGFGCNLTQCCLNGGACN
TLDGTCTCAPGWRGEKCELPQDGT~~Y~~GLNCAERCDCSHADGCHPTTGHCRC~~LP~~GWSGV

HCDSVCAEGRWGPNC SLPCYCKNGASCSPDDGICECAPGFRGTTTCQRICSPGFYGHRC
 SQTCPQCVHSSGPCHHITGLCDCLPGFTGALCNEVCPSGRFGKNCAGICTCTNNGTCN
 PIDRSCQCYPGWIGSDCSQPCPPAHWGPNCIHTCNCHNGAFCSAYDGECKCTPGWTGL
 YCTQRCPLGFYKDCALICQCQNGADCDHISGQCTCRTGFMGRHCEQKCPSTYGYGC
 RQICDCLNNSTCDHITGTCTCYCSPGWKGARCDQAGVIVGNLNSLSRTSTALPADSYQI
 GAIAGIIILVLVFLFLLALFIIYRHKQKGKSSMPAVTYTPAMRVVNADYTIISGTLPH
 SNGGNANSHYFTNPSYHTLTQCATS PHVNNRDRMTVTKSKNNQLFVN LKNVNPGRGP
 VGDCTGTLPADWKHGGYLNELGAFGLDRSYMGS LKDLGKNSEYNSSNCSLSSSENPY
 ATIKDPPVLIPKSSECGYVEMKSPARRDSPYAEINNSTSANRNVYEVEPTVSVVQGVF
 SNNGRLSQDPYDLPKNSHIPCHYDLLPVRDSSSSPKQEDSGGSSSSNSSSSSE"

ORIGIN

Query Match 100.0%; Score 3421.4; DB 9; Length 7522;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT	60
Db	204	ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT	263
Qy	61	GGGACAGCATCACCTCTGAATCTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC	120
Db	264	GGGACAGCATCACCTCTGAATCTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC	323
Qy	121	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	180
Db	324	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	383
Qy	181	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	240
Db	384	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	443
Qy	241	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	300
Db	444	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	503
Qy	301	AGCGGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	360
Db	504	AGCGGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	563
Qy	361	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	420
Db	564	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	623
Qy	421	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	480
Db	624	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	683
Qy	481	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	540
Db	684	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	743
Qy	541	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	600
Db	744	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	803
Qy	601	GCCACCTGCGACCACGTCACGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTC	660

Db	804	 GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTC	863
Qy	661	TGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Db	864	 TGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	923
Qy	721	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Db	924	 CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	983
Qy	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Db	984	 GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	1043
Qy	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Db	1044	 TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	1103
Qy	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960
Db	1104	 GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	1163
Qy	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1020
Db	1164	 GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1223
Qy	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGCTCTAC	1080
Db	1224	 TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGCTCTAC	1283
Qy	1081	GGCATCAAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTACCCCC	1140
Db	1284	 GGCATCAAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTACCCCC	1343
Qy	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Db	1344	 ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1403
Qy	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Db	1404	 TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1463
Qy	1261	TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Db	1464	 TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1523
Qy	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAAT	1380
Db	1524	 ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAAT	1583
Qy	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGACTTGCAAGGCAGGCTGGCACGGGGTG	1440
Db	1584	 GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGACTTGCAAGGCAGGCTGGCACGGGGTG	1643
Qy	1441	GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTTAACATGCCAG	1500

Db	1644	GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1703
Qy	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1560
Db	1704	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1763
Qy	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1620
Db	1764	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1823
Qy	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1680
Db	1824	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1883
Qy	1681	CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1740
Db	1884	CCCGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1943
Qy	1741	TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGC	1800
Db	1944	TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGC	2003
Qy	1801	GAGTGTGCACCAAGGCTTCCGAGGCACCACCTTGTGACAGGATCTGCTCCCCTGGTTTTAT	1860
Db	2004	GAGTGTGCACCAAGGCTTCCGAGGCACCACCTTGTGACAGGATCTGCTCCCCTGGTTTTAT	2063
Qy	1861	GGGCATCGCTGCAGCCAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCAC	1920
Db	2064	GGGCATCGCTGCAGCCAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCAC	2123
Qy	1921	ATCACCGGCCTGTGTGACTGCTTGCCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT	1980
Db	2124	ATCACCGGCCTGTGTGACTGCTTGCCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT	2183
Qy	1981	CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC	2040
Db	2184	CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC	2243
Qy	2041	TGTAACCCCATTTGACAGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTCT	2100
Db	2244	TGTAACCCCATTTGACAGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTCT	2303
Qy	2101	CAACCATGTCCACCTGCCCACTGGGGCCCAACTGCATCCACACGTGCAACTGCCATAAT	2160
Db	2304	CAACCATGTCCACCTGCCCACTGGGGCCCAACTGCATCCACACGTGCAACTGCCATAAT	2363
Qy	2161	GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC	2220
Db	2364	GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC	2423
Qy	2221	TACTGCACTCAGAGATGTCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAA	2280
Db	2424	TACTGCACTCAGAGATGTCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAA	2483
Qy	2281	TGTCAAAACGGAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTC	2340
Db	2484	TGTCAAAACGGAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTC	2543

Qy	2341	ATGGGACGGCACTGTGAGCAGAAAGTGCCCTTCAGGAACATATGGCTATGGCTGTCGCCAG	2400
Db	2544	ATGGGACGGCACTGTGAGCAGAAAGTGCCCTTCAGGAACATATGGCTATGGCTGTCGCCAG	2603
Qy	2401	ATATGTGATTGTCTGAACAACCTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC	2460
Db	2604	ATATGTGATTGTCTGAACAACCTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC	2663
Qy	2461	CCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGTTATCATAGTTGGAAATCTGAAC	2520
Db	2664	CCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGTTATCATAGTTGGAAATCTGAAC	2723
Qy	2521	AGCTTAAGCCGAACCAGTACTGCTCTCCCTGCTGATTCTACCAGATCGGGGCCATTGCA	2580
Db	2724	AGCTTAAGCCGAACCAGTACTGCTCTCCCTGCTGATTCTACCAGATCGGGGCCATTGCA	2783
Qy	2581	GGCATCATCATTCTTGTCTAGTTGTTCTCTTCTACTGGCATTGTTCAATTATTTATAGA	2640
Db	2784	GGCATCATCATTCTTGTCTAGTTGTTCTCTTCTACTGGCATTGTTCAATTATTTATAGA	2843
Qy	2641	CACAAGCAGAAGGGAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2700
Db	2844	CACAAGCAGAAGGGAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2903
Qy	2701	GTCGTCAATGCAGATTATACCATTTAGGAACCCTTCCTCACAGCAATGGTGGAAACGCT	2760
Db	2904	GTCGTCAATGCAGATTATACCATTTAGGAACCCTTCCTCACAGCAATGGTGGAAACGCT	2963
Qy	2761	AATAGCCACTACTTCACCAATCCCAGTTACCACACGCTCACCCAGTGTCACACATCCCCCT	2820
Db	2964	AATAGCCACTACTTCACCAATCCCAGTTACCACACGCTCACCCAGTGTCACACATCCCCCT	3023
Qy	2821	CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	2880
Db	3024	CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	3083
Qy	2881	AATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	2940
Db	3084	AATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	3143
Qy	2941	CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3000
Db	3144	CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3203
Qy	3001	AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAAGAATTCTGAATATAATTCAAGTAAC	3060
Db	3204	AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAAGAATTCTGAATATAATTCAAGTAAC	3263
Qy	3061	TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3120
Db	3264	TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3323
Qy	3121	CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCA	3180
Db	3324	CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCA	3383

Qy 3181 TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA 3240
 |||
 Db 3384 TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA 3443
 Qy 3241 GTGAGTGTGTCCAAGGAGTATTCAGCAATAATGGGCGTCTCTCCAGGATCCATATGAC 3300
 |||
 Db 3444 GTGAGTGTGTCCAAGGAGTATTCAGCAATAATGGGCGTCTCTCCAGGATCCATATGAC 3503
 Qy 3301 CTCCCAAAGAACAGTCACATCCCTTGTCAATTATGACCTGCTGCCAGTCCGAGACAGTTCA 3360
 |||
 Db 3504 CTCCCAAAGAACAGTCACATCCCTTGTCAATTATGACCTGCTGCCAGTCCGAGACAGTTCA 3563
 Qy 3361 TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGCAGTGAA 3420
 |||
 Db 3564 TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGCAGTGAA 3623
 Qy 3421 TGA 3423
 |||
 Db 3624 TGA 3626

RESULT 3

AK123568

LOCUS AK123568 2952 bp mRNA linear PRI 09-SEP-2003
 DEFINITION Homo sapiens cDNA FLJ41574 fis, clone CTONG2010116, weakly similar to Rattus norvegicus mRNA for MEGF6.

ACCESSION AK123568

VERSION AK123568.1 GI:34529147

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M., Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2952)

AUTHORS Isogai,T. and Yamamoto,J.

TITLE Direct Submission

JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan

COMMENT (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and


```

RAB; annotation: HRI and RAB.
FEATURES
    source          Location/Qualifiers
                    1. .2952
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="CTONG2010116"
                    /tissue_type="tongue, tumor tissue"
                    /clone_lib="CTONG2"
                    /note="cloning vector: pME18SFL3"

```

ORIGIN

```

Query Match          66.5%;   Score 2274.8;   DB 9;   Length 2952;
Best Local Similarity 99.9%;   Pred. No. 0;
Matches 2276;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

```

```

Qy      1 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 60
        |||
Db      478 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 537

Qy      61 GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC 120
        |||
Db      538 GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC 597

Qy      121 TCAGTGAAGTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC 180
        |||
Db      598 TCAGTGAAGTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC 657

Qy      181 ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT 240
        |||
Db      658 ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT 717

Qy      241 CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA 300
        |||
Db      718 CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA 777

Qy      301 AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT 360
        |||
Db      778 AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT 837

Qy      361 CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT 420
        |||
Db      838 CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT 897

Qy      421 GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC 480
        |||
Db      898 GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC 957

Qy      481 AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC 540
        |||
Db      958 AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC 1017

Qy      541 CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA 600
        |||
Db      1018 CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA 1077

Qy      601 GCCACCTGCGACCACGTCACGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTC 660

```

Db	1078	 GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTC	1137
Qy	661	TGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Db	1138	 TGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	1197
Qy	721	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Db	1198	 CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	1257
Qy	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Db	1258	 GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	1317
Qy	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Db	1318	 TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	1377
Qy	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960
Db	1378	 GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	1437
Qy	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1020
Db	1438	 GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1497
Qy	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCGTGTGCCTGAGGGGCTCTAC	1080
Db	1498	 TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCGTGTGCCTGAGGGGCTCTAC	1557
Qy	1081	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTCACCCC	1140
Db	1558	 GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTCACCCC	1617
Qy	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Db	1618	 ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1677
Qy	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Db	1678	 TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1737
Qy	1261	TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Db	1738	 TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1797
Qy	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1380
Db	1798	 ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1857
Qy	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGACTTGCAAGGCAGGCTGGCACGGGGTG	1440
Db	1858	 GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGACTTGCAAGGCAGGCTGGCACGGGGTG	1917
Qy	1441	GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1500

Db 1918 GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG 1977

Qy 1501 TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG 1560
 |||

Db 1978 TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG 2037

Qy 1561 CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG 1620
 |||

Db 2038 CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG 2097

Qy 1621 CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC 1680
 |||

Db 2098 CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC 2157

Qy 1681 CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC 1740
 ||

Db 2158 CCCGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC 2217

Qy 1741 TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCTGATGATGGCATCTGC 1800
 |||

Db 2218 TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCTGATGATGGCATCTGC 2277

Qy 1801 GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCTGGTTTTTAT 1860
 |||

Db 2278 GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCTGGTTTTTAT 2337

Qy 1861 GGGCATCGCTGCAGCCAGACATGCCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCAC 1920
 |||

Db 2338 GGGCATCGCTGCAGCCAGACATGCCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCAC 2397

Qy 1921 ATCACC GGCCCTGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT 1980
 |||

Db 2398 ATCACC GGCCCTGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT 2457

Qy 1981 CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC 2040
 |||

Db 2458 CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC 2517

Qy 2041 TGTAACCCCATTTGACAGATCTTGTGAGTGTACCCCGTTGGATTGGCAGTGACTGCTCT 2100
 |||

Db 2518 TGTAACCCCATTTGACAGATCTTGTGAGTGTACCCCGTTGGATTGGCAGTGACTGCTCT 2577

Qy 2101 CAACCATGTCCACCTGCCCCTGGGGCCCAACTGCATCCACACGTGCAACTGCCATAAT 2160
 |||

Db 2578 CAACCATGTCCACCTGCCCCTGGGGCCCAACTGCATCCACACGTGCAACTGCCATAAT 2637

Qy 2161 GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC 2220
 |||

Db 2638 GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC 2697

Qy 2221 TACTGCACTCAGAGATGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCC 2278
 |||

Db 2698 TACTGCACTCAGAGATGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGAC 2755

LOCUS BC020198 2267 bp mRNA linear PRI 16-SEP-2003
 DEFINITION Homo sapiens MEGF10 protein, mRNA (cDNA clone IMAGE:4904255), complete cds.
 ACCESSION BC020198
 VERSION BC020198.1 GI:18044365
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2267)
 AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 2267)
 AUTHORS Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 40 Row: 0 Column: 4

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14192942

This clone has the following problem: The cds is short compared to the longest cds in the locus.

```
FEATURES             Location/Qualifiers
     source            1. .2267
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:4904255"
                        /tissue_type="Muscle, rhabdomyosarcoma"
                        /clone_lib="NIH_MGC_17"
                        /lab_host="DH10B-R"
                        /note="Vector: pOTB7"
     gene              1. .2267
                        /gene="MEGF10"
                        /note="synonym: KIAA1780"
                        /db_xref="LocusID:84466"
     CDS                280. .1983
                        /codon_start=1
                        /product="MEGF10 protein"
                        /protein_id="AAH20198.1"
                        /db_xref="GI:18044366"
                        /db_xref="LocusID:84466"
                        /translation="MVISLNSCLSFICLLLCHWIGTASPLNLEDPNVCSHWESYSVTV
QESYPHPFDQIYYTSCTDILNWFKCTRHRVSYRTAYRHGEKTMYRRKSQCCPGFYESG
EMCVPHCADKCVHGRCIAPNTCQCEPGWGGTNCSSACDGDHWGPHCTSRCQCKNGALC
NPITGACHCAAGFRGWRCEDRCEQGTYGNDCHQRCQCQNGATCDHVTGECRCPPGYTG
AFCEDLCPPGKHGPQCEQRCPCQNGGVCHHVTGECSCPSGWMGTVCQGPCPEGRFGKN
CSQECQCHNGGTCDAAATQCHCSPGYTGERCQDECPVGTYGVLCAETCQCVNGGKCYH
VSGACLCEAGFAGERCEARLCPEGLYGIKCDKRCPCHLENTHSCHPMSGECACKPGWS
GLYCNETCSPGFYGEACQQICSCQNGADCDSVTGKCTCAPGFKIDCSTPCPLGTYGI
NCSSRCGCKNDAVCSPVDGSCTCKAGWHGVDCSIRCPSGTWGFGCNLTCQCLNGGACN
TLDGTCTCAPGWRGEKCELPCQDGTYGLNCAERCDCSHADGCHPTTGHCRCCLPGWSGL
F"
```

ORIGIN

Query Match 49.4%; Score 1692.4; DB 9; Length 2267;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1693; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 60
        |||
Db      280 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 339

Qy      61 GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC 120
        |||
Db      340 GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC 399

Qy      121 TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC 180
        |||
Db      400 TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC 459
```

Qy	181	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	240
Db	460	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	519
Qy	241	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	300
Db	520	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	579
Qy	301	AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	360
Db	580	AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	639
Qy	361	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	420
Db	640	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	699
Qy	421	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	480
Db	700	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	759
Qy	481	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	540
Db	760	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	819
Qy	541	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	600
Db	820	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	879
Qy	601	GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCACCAGGATACACGGAGCCTTC	660
Db	880	GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCACCAGGATACACGGAGCCTTC	939
Qy	661	TGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Db	940	TGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	999
Qy	721	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Db	1000	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	1059
Qy	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Db	1060	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	1119
Qy	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Db	1120	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	1179
Qy	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960
Db	1180	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	1239
Qy	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTTACCACGTGAGCGGCGCATGCCTC	1020
Db	1240	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTTACCACGTGAGCGGCGCATGCCTC	1299

Qy	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGGCTCTAC	1080
Db	1300	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGGCTCTAC	1359
Qy	1081	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTCACCCC	1140
Db	1360	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTCACCCC	1419
Qy	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Db	1420	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1479
Qy	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Db	1480	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1539
Qy	1261	TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Db	1540	TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1599
Qy	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1380
Db	1600	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1659
Qy	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTG	1440
Db	1660	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTG	1719
Qy	1441	GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1500
Db	1720	GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1779
Qy	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1560
Db	1780	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1839
Qy	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1620
Db	1840	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1899
Qy	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1680
Db	1900	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1959
Qy	1681	CCGGGATGGTCAGG	1694
Db	1960	CCCGGATGGTCAGG	1973

RESULT 5

AR217554

LOCUS AR217554 1448 bp DNA linear PAT 25-SEP-2002

DEFINITION Sequence 9 from patent US 6416974.

ACCESSION AR217554

VERSION AR217554.1 GI:23317353

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 1448)
 AUTHORS Holtzman,D.A. and Goodearl,A.D.J.
 TITLE Tango 71 nucleic acids
 JOURNAL Patent: US 6416974-A 9 09-JUL-2002;
 FEATURES Location/Qualifiers
 source 1. .1448
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 41.7%; Score 1425.8; DB 6; Length 1448;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1427; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1216	GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT	1275
Db	18	GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT	77
Qy	1276	GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG	1335
Db	78	GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG	137
Qy	1336	GGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT	1395
Db	138	GGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT	197
Qy	1396	CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGA	1455
Db	198	CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGA	257
Qy	1456	TGTCCCACTGGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGA	1515
Db	258	TGTCCCACTGGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGA	317
Qy	1516	GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC	1575
Db	318	GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC	377
Qy	1576	GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC	1635
Db	378	GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC	437
Qy	1636	CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT	1695
Db	438	CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT	497
Qy	1696	GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC	1755
Db	498	GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC	557
Qy	1756	TACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC	1815
Db	558	TACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC	617
Qy	1816	TTCCGAGGCACCACTTGTGACAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCTGCAGC	1875

Db	618	TTCCGAGGCACCACTTGTCTCAGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCTGCAGC	677
Qy	1876	CAGACATGCCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGT	1935
Db	678	CAGACATGCCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGT	737
Qy	1936	GACTGCTTGCCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT	1995
Db	738	GACTGCTTGCCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT	797
Qy	1996	GGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTGAC	2055
Db	798	GGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTGAC	857
Qy	2056	AGATCTTGTCTAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTCTCAACCATGTCCACCT	2115
Db	858	AGATCTTGTCTAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTCTCAACCATGTCCACCT	917
Qy	2116	GCCCACTGGGGCCCAAACCTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC	2175
Db	918	GCCCACTGGGGCCCAAACCTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC	977
Qy	2176	GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA	2235
Db	978	GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA	1037
Qy	2236	TGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCT	2295
Db	1038	TGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCT	1097
Qy	2296	GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGCACTGT	2355
Db	1098	GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGCACTGT	1157
Qy	2356	GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGTCTG	2415
Db	1158	GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGTCTG	1217
Qy	2416	AACAACCTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA	2475
Db	1218	AACAACCTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA	1277
Qy	2476	GCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC	2535
Db	1278	GCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC	1337
Qy	2536	AGTACTGCTCTCCCTGCTGATTCCCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTT	2595
Db	1338	AGTACTGCTCTCCCTGCTGATTCCCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTT	1397
Qy	2596	GTCCTAGTTGTTCTCTTCTTCTACTGGCATTGTTTATTATTTATAGACACA	2644
Db	1398	GTCCTAGTTGTTCTCTTCTTCTACTGGCATTGTTTATTATTTATAGACACA	1446

BD080296
 LOCUS BD080296 1448 bp DNA linear PAT 27-AUG-2002
 DEFINITION Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 nucleic acid molecules and polypeptides.
 ACCESSION BD080296
 VERSION BD080296.1 GI:22625899
 KEYWORDS JP 2001512681-A/5.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1448)
 AUTHORS Holtzman,D.A. and Goodearl,A.D.J.
 TITLE Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 nucleic acid molecules and polypeptides
 JOURNAL Patent: JP 2001512681-A 5 28-AUG-2001;
 MILLENNIUM PHARMACEUTICALS INC
 COMMENT OS Homo sapiens (human)
 PN JP 2001512681-A/5
 PD 28-AUG-2001
 PF 06-AUG-1998 JP 2000506335
 PR 06-AUG-1997 US 60/054966,05-SEP-1997 US 60/058108 PI
 DOUGLAS A HOLTZMAN,ANDREW D J GOODEARL
 PC C12N15/09,C07K14/47,C07K16/18,C12N5/10,C12P21/02,C12Q1/68, PC
 G01N33/15,
 PC G01N33/50,G01N33/53//C12P21/08,C12N15/00,C12N5/00 CC
 Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 nucleic acid
 acid
 CC molecules and polypeptides
 FH Key Location/Qualifiers
 FT source 1. .1448
 FT /organism='Homo sapiens (human)'.
 FEATURES Location/Qualifiers
 source 1. .1448
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 41.7%; Score 1425.8; DB 6; Length 1448;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1427; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1216 GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT 1275
 |||
 Db 18 GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT 77

Qy 1276 GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG 1335
 |||
 Db 78 GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG 137

Qy 1336 GGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT 1395
 |||
 Db 138 GGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT 197

Qy 1396 CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGGCACGGGGTGGACTGCTCCATCAGA 1455
 |||

Db 198 CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGA 257

Qy 1456 TGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGA 1515
 |||

Db 258 TGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGA 317

Qy 1516 GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC 1575
 |||

Db 318 GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC 377

Qy 1576 GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC 1635
 |||

Db 378 GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC 437

Qy 1636 CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT 1695
 |||

Db 438 CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT 497

Qy 1696 GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC 1755
 |||

Db 498 GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC 557

Qy 1756 TACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC 1815
 |||

Db 558 TACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC 617

Qy 1816 TTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCTGCAGC 1875
 |||

Db 618 TTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCTGCAGC 677

Qy 1876 CAGACATGCCACAGTGCCTTACAGCAGCGGGCCCTGCCACCACATCACCAGGCCTGTGT 1935
 |||

Db 678 CAGACATGCCACAGTGCCTTACAGCAGCGGGCCCTGCCACCACATCACCAGGCCTGTGT 737

Qy 1936 GACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT 1995
 |||

Db 738 GACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT 797

Qy 1996 GGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTGAC 2055
 |||

Db 798 GGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTGAC 857

Qy 2056 AGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTCTCAACCATGTCCACCT 2115
 |||

Db 858 AGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTCTCAACCATGTCCACCT 917

Qy 2116 GGGCACTGGGGCCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC 2175
 |||

Db 918 GGGCACTGGGGCCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC 977

Qy 2176 GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA 2235
 |||

Db 978 GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA 1037

Qy 2236 TGTCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCT 2295
 |||

Db 1038 TGTCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCT 1097

```

Qy      2296 GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGCACTGT 2355
          |||
Db      1098 GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGCACTGT 1157

Qy      2356 GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTCGCCAGATATGTGATTGTCTG 2415
          |||
Db      1158 GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTCGCCAGATATGTGATTGTCTG 1217

Qy      2416 AACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA 2475
          |||
Db      1218 AACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA 1277

Qy      2476 GCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC 2535
          |||
Db      1278 GCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC 1337

Qy      2536 AGTACTGCTCTCCCTGCTGATTCCCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTT 2595
          |||
Db      1338 AGTACTGCTCTCCCTGCTGATTCCCTACCAATCGGGGCCATTGCAGGCATCATCATTCTT 1397

Qy      2596 GTCCTAGTTGTTCTCTTCCTACTGGCATTGTTTCATTATTTATAGACACA 2644
          |||
Db      1398 GTCCTAGTTGTTCTCTTCCTACTGGCATTGTTTCATTATTTATAGACACA 1446

```

RESULT 7

AB058677

LOCUS AB058677 5702 bp mRNA linear PRI 10-MAY-2001

DEFINITION Homo sapiens mRNA for MEGF11 protein (KIAA1781), complete cds.

ACCESSION AB058677

VERSION AB058677.1 GI:14017778

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)

AUTHORS Nagase,T., Nakayama,M., Nakajima,D., Kikuno,R. and Ohara,O.

TITLE Prediction of the coding sequences of unidentified human genes. XX.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro

JOURNAL DNA Res. 8 (2), 85-95 (2001)

MEDLINE 21245130

PUBMED 11347906

REFERENCE 2 (bases 1 to 5702)

AUTHORS Nakayama,M., Nagase,T., Nakajima,D., Kikuno,R. and Ohara,O.

TITLE Direct Submission

JOURNAL Submitted (27-MAR-2001) Manabu Nakayama, Kazusa DNA Research
Institute, Department of Human Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:nmanabu@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3915,
Fax:81-438-52-3914)

FEATURES Location/Qualifiers

source 1. .5702

/organism="Homo sapiens"

/mol_type="mRNA"

```

/db_xref="taxon:9606"
/clone="fg06971"
/tissue_type="brain"
/dev_stage="fetal"
/note="vector:pBluescript II SK plus"
gene 1. .5702
/ gene="MEGF11"
CDS 160. .3069
/ gene="MEGF11"
/ note="KIAA1781 protein
Start codon is not confirmed. fg06971 cDNA clone for
KIAA1781 has a 119bp insertion after the position 2642 of
the sequence of KIAA1781
gene encoding protein with multiple EGF-like-domains"
/ codon_start=1
/ product="MEGF11 protein (KIAA1781)"
/ protein_id="BAB47410.1"
/ db_xref="GI:14017779"
/ translation="MHTPSIRSITHDAQTSSTGSSAPGTALCTEECVHGRCVSPDTCH
CEPGWGGPDCSSGCDSDHWGPHCSNRCQCQNGALCNPITGACVCAAGFRGWRCEELCA
PGTHGKGCQLPCQCRHGASCDPRAGECLCAPGYTGVCYCEELCPPGSHGAHCELCRPCQ
NGGTCHHITGECACPPGWTGAVCAQPCPPGTFGQNCSDQDCPCHHGGQCDHVTGQCHCT
AGYMGDRQCQEECPFGSFGFQCSQRCDCCHNGGQCSPTTGACECEPGYKGFRCQERLCPE
GLHGPGCTLPCPCDADNTISCHPVTGACTCQPGWSGHHCNESCPVGYGDCGQLPCTC
QNGADCHSITGGCTCAPGFMGEVCAVSCAAGTYGPNCSISCSNNGGTCSPVDGSCCTC
KEGWQGLDCTLPCPSGTWGLNCNESCTCANGAACSPIDGSCSCTPGWLGDTCCLPCPD
GTFGLNCSEHDCSHADGCDPVTGHCCCLAGWTGIRCDSTCPPGRWGPNCVSVSCSEN
GGSCSPEDGSCCAPGFRGPLCQRICPPGFYGHGCAQPCPLCVHSSRPCHHISGICEC
LPGFSGALCNQVCAGGYFGQDCAQLCSCANNGTCSPIDGSCQCFPGWIGKDCSQACPP
GFWGPACFHACSCHNGASCSAEDGACHCTPGWTGLFCTQRCPAAFFGKDCGRVCQCQN
GASCDHISGKCTCRTGFTGQHCEQRCAPGTFGYGCQQLCECMNNSTCDHVTGTCTCYCSP
GFKGIRCDQAALMMEELNPYTKISPALGAERHVSVAVTGIMLLFFIVVLLGLFAWHR
RRQKEKGRDLAPRVSYTPAMRMTSTDYSLSGACGMDRRQNTYIMDKGFKDYMKEVCS
SSTCSLNSSNPYATIKDPPILTCKLPESYVEMKSPVHMGSPYTDVPSLSTSNKNIY
EVEPTVSVVQEGCGHNSSYIQNAYDLPRNSHIPGHYDLLPVRQSPANGPSQDKQS"

```

ORIGIN

```

Query Match          35.1%;  Score 1200.6;  DB 9;  Length 5702;
Best Local Similarity 69.2%;  Pred. No. 0;
Matches 1670;  Conservative 0;  Mismatches 739;  Indels 6;  Gaps 2;

Qy 318 CCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCTCCAAACACCTGTTCAGTG 377
    | ||| ||| | || | |||| || || |||| | || | || |||| || ||
Db 234 CGCCCTGTGTACGGAGGAGTGTGTGCACGGCCGCTGCGTTTCCCCGGACACCTGCCACTG 293

Qy 378 TGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCCTGCGATGGTGATCACTGGGGTCC 437
    ||||| ||||| ||||| || ||||| || ||||| || || ||||| || ||
Db 294 CGAGCCTGGCTGGGGAGGGCCCCACTGCTCCAGCGGCTGCGACAGCGACCACTGGGGGCC 353

Qy 438 CCACTGCACCAAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGCAACCCCATCACCGGGGC 497
    ||||| || ||||| ||||| || || || || || || ||||| || || ||
Db 354 CCACTGCAGCAACCGGTGCCAGTGCCAGAACGGCGCCCTGTGTAACCCCATCACAGGCGC 413

Qy 498 TTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGCAGGGCAC 557
    ||| || || || ||||| || ||||| ||||| || || || |||||
Db 414 CTGCGTGTGCGCCGCCGGCTTCCGTGGATGGCGCTGCGAGGAGCTCTGCGCGCCTGGCAC 473

```

[illegible]

Db	1314	AGTAGATGGCTCCTGTACCTGCAAGGAAGGGTGGCAGGGCCTGGACTGCACCCTGCCATG	1373
Qy	1458	TCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGAGC	1517
Db	1374	TCCCAGTGGGACGTGGGGCCTGAACTGCAACGAGAGCTGCACCTGTGCCAATGGGGCAGC	1433
Qy	1518	CTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGCGA	1577
Db	1434	CTGCAGCCCCATAGACGGCTCCTGCTCCTGCACTCCTGGCTGGCTGGGAGACACCTGTGA	1493
Qy	1578	ACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCCACTGCAGCCA	1637
Db	1494	GCTGCCTTGCCCGGATGGCACATTTGGGCTGAACTGCAGTGAACACTGTGACTGCAGCCA	1553
Qy	1638	CGCAGATGGCTGCCACCCTACCACGGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGTGT	1697
Db	1554	TGCTGATGGATGTGACCCCGTCACAGGCCACTGCTGCTGCCTGGCCGGATGGACAGGCAT	1613
Qy	1698	CCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGCTA	1757
Db	1614	CCGCTGTGACAGCACGTGTCCACCTGGCCGCTGGGGCCCCAACTGCTCTGTCTCCTGCAG	1673
Qy	1758	CTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGCTT	1817
Db	1674	CTGTGAGAATGGAGGCTCCTGCTCCCCAGAGGATGGGAGCTGCGAGTGTGCCCTTGCTT	1733
Qy	1818	CCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCTGCAGCCA	1877
Db	1734	CCGAGGACCCTTATGCCAGAGAATCTGCCCCCTGGGTTCATGGCCACGGCTGCGCCCA	1793
Qy	1878	GACATGCCCACAGTGC GTTACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGTGA	1937
Db	1794	GCCATGCCCCCTCTGCGTGACAGCAGCAGGCCCTGCCACCACATCAGCGGCATCTGTGA	1853
Qy	1938	CTGCTTGCTGGCTTACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTTGG	1997
Db	1854	GTGCCTCCCAGGATTCTCTGGAGCTCTCTGCAACCAAGTGTGTGCTGGAGGATACTTTGG	1913
Qy	1998	GAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTTGACAG	2057
Db	1914	GCAGGACTGTGCCAGCTCTGCTCCTGTGCCAACAAACGGGACCTGCAGCCCTATCGATGG	1973
Qy	2058	ATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGA CTGCTCAACCATGTCCACCTGC	2117
Db	1974	CTCCTGCCAGTGCTTTCTGGATGGATTGGCAAGGACTGCTCACAGGCTTGCCACCCCGG	2033
Qy	2118	CCACTGGGGCCCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGCGC	2177
Db	2034	GTTCTGGGGCCCCGCCTGCTTCCACGCATGCAGCTGCCACAACGGGGCGAGCTGCAGCGC	2093
Qy	2178	CTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGATG	2237
Db	2094	CGAGGACGGGGCCTGCCACTGCACCCCTGGCTGGACTGGACTCTTCTGCACACAGCGCTG	2153
Qy	2238	TCCTCTAGGGTTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACGGAGCTGA	2297

Db 2154 CCCAGCAGCATT TTTTGGGAAGGACTGTGGGCGCGTATGCCAGTGT CAGAATGGCGCCAG 2213

Qy 2298 CTGCGACCACATT TCTGGGCAGTGTACTTGCCGCACTGGATT CATGGGACGGCACTGTGA 2357
 ||| ||||| ||| |||| || ||||| || |||| || | |||||

Db 2214 CTGTGACCACATCAGTGGCAAGTGCACCTGCCGCACAGGCTTACC GGGCAACACTGTGA 2273

Qy 2358 GCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGT CGCCAGATATGTGATTGTCTGAA 2417
 ||||| || | ||||| | ||||| |||| ||| ||||| || ||||

Db 2274 GCAGAGATGTGCCCCAGGAACCTTTGGCTATGGGTGTCAGCAGCTATGTGAGTGCATGAA 2333

Qy 2418 CAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGAGC 2477
 ||||| |||| |||| || ||||| ||||| || | || |||

Db 2334 CAACTCCACCTGTGACCATGTCACCGGCACCTGTTACTGCAGCCCTGGCTTCAAAGGAAT 2393

Qy 2478 GAGATGTGATCAAGCTGGTGT TATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACCAG 2537
 || |||| ||||| || || || | ||||| || || | |||

Db 2394 CAGGTGTGACCAAGCTGCCCTCATGATGGAGG---AGCTGAATCCCTACACCAAGATCAG 2450

Qy 2538 TACTGCTCTCCCTGCTGATTCCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTTGT 2597
 | || || ||| || || || | || || | ||||| || || |

Db 2451 CCCAGCACTGGGTGCAGAGCGGCACCTCGGTGGGTGCTGT CACAGGCATCATGCTCCTGTT 2510

Qy 2598 CCTAGTTGTTCTCTTCTACTGGCATTGTT CATTATTTATAGACACAAGCAGAAGGGAAA 2657
 | | || | || |||| | || || || | |||| | ||

Db 2511 ATTCTTCATTGTGGTGCTGCTGGGCCTATTTGCCTGGCATCGGCGGCGGCAGAAAGAGAA 2570

Qy 2658 GGAATCAAGCATG---CCAGCAGTTACCTACACCCCTGCTATGAGGGTCGTCAATGCAGA 2714
 || | | || || || ||||| |||| ||||| | || | ||

Db 2571 GGGCCGAGACCTGGCTCCCCGTGTCTCCTACACACCTGCCATGAGGATGACCAGCACC GA 2630

Qy 2715 TTATACCATTTCAGG 2729
 || || | ||||

Db 2631 CTA CTCTCCCTCTCAGG 2645

RESULT 8

AK122555

LOCUS AK122555 5278 bp mRNA linear ROD 15-MAR-2003

DEFINITION Mus musculus mRNA for mKIAA1781 protein.

ACCESSION AK122555

VERSION AK122555.1 GI:28972841

KEYWORDS FLI_CDNA.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S., Nakajima,D., Nagase,T., Ohara,O. and Koga,H.

TITLE Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries

JOURNAL DNA Res. 10, 35-48 (2003)

REFERENCE 2 (bases 1 to 5278)

AUTHORS Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.

TITLE Direct Submission

JOURNAL Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)

COMMENT The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing: Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.

FEATURES Location/Qualifiers

source 1. .5278
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mbg06605"
/tissue_type="brain"
/dev_stage="adult"
/note="vector:modified pBC SK+"

gene 1. .5278
/gene="mKIAA1781"

CDS <1. .2766
/gene="mKIAA1781"
/note="CDS is predicted by in silico analysis. Start codon
is not identified."
/codon_start=1
/evidence=not_experimental
/product="mKIAA1781 protein"
/protein_id="BAC65837.1"
/db_xref="GI:28972842"
/translation="CPPGSHGAHCELRCPCQNGGTCHHITGECACPPGWTGAVCAQPC
PPGTFGQNCSDCPCHHGGQCDHVTGQCHCTAGYMGDRCQECPFGTFGFLCSQRCD
HNGGQCSPATGACECEPGYKGPSCQERLCPEGLHGPCTLPCCDENTISCHPVTGA
CTCQPGWSGHYCNESCPAGYYGNGCQLPCTCQNGADCHSITGSCTCAPGFMGEVCAVP
CAAGTYGPNCSVCSCSNGGTCTSPVDGSCTCREGWQGLDCSLPCPSGTWGLNCNETCI
CANGAACSPFDGSCACTPGWLGDSCLEPCPDGTFGLNCSEHCDCSHADGCDPVTGHCC
CLAGWTGIRCDSTCPPGRWGPNCVSCSCENGSCSPEDGSCECAPGFRGPLCQRICP
PGFYGHGCAQPCPLCVHSRGPCHHISGICECLPGFSGALCNQVCAGGHFGQDCAQLCS
CANNGTCSPIDGSCQCFPGWIGKDCSQACPSGFWGSACFHTCSCHNGASCSAEDGACH
CTPGWTGLFCTQRCPSAFFGKDCGHICQCQNGASCDHITGKCTCRTGFSGRHCEQRCA
PGTFGYGCQQLCECMNNATCDHVTGTCTCYCSPGFKGIRCDQAALMMDELNPYTKISPAL
GAERHSVGAVTGIVLLFLVVVLLGLFAWRRRRQKEKGRDLAPRVSYTPAMRMTSTDY
SLSDLQSSSSHAQCFSNASYHTLACGGPATSQASTLDRNSPTKLSNKS LDRDTAGWTP
YSYVNVLD SHFQISALEARYPPEDFYIELRHL SRHAEPHSPGTCGMDRRQNTYIMDKG
FKDYMKESVCSSSTCSLNSEN PYATIKDPPILTCKLPESYVEMKSPVHLGSPYTDV
PSLSTSNKNIYEVEPTVSVVQEGRGHNSSYIQNPYDLPKNSHIPGHYDLLPVRQSPA
HPFQEKQ"

ORIGIN

Query Match 30.4%; Score 1040.8; DB 10; Length 5278;
Best Local Similarity 63.3%; Pred. No. 1.6e-310;
Matches 1736; Conservative 0; Mismatches 952; Indels 54; Gaps 7;

Qy 673 TGTCTCTCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAATGGAGGA 732
|| || ||||| | ||||| | || ||||| | || ||||| || || ||||| ||

Db 1 TGCCCCCTGGGAGCCATGGAGCTCACTGTGAGCTGCGCTGCCCTGCCAGAATGGAGGC 60

Qy 733 GTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGTGT 792
|| || ||| ||||| ||||| | ||||| | ||||| || ||||| ||

Db		61	ACCTGCCACCACATCACTGGCGAATTGTCCTGCCTCCAGGCTGGACGGGAGCAGTGTGT	120
Qy		793	GGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCAAGAATGCCAGTGCCAT	852
Db		121	GCCCAGCCCTGCCCTCCAGGGACCTTTGGCCAGAACTGTAGCCAGGACTGTCCCTGCCAC	180
Qy		853	AATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAGGG	912
Db		181	CATGGAGGCCAGTGTGACCATGTGACTGGACAATGCCACTGTACAGCTGGATACATGGGG	240
Qy		913	GAACGGTGCCAGGATGAGTGTCTGTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGC	972
Db		241	GACAGGTGTCAAGAAGAATGTCCCTTTGGAACGTTTCGGTTTCTGTGCTCTAACGCTGT	300
Qy		973	CAGTGTGTCAACGGAGGGAAAGTGTACACAGTGAAGCGCGCATGCCCTCTGTGAAGCAGGC	1032
Db		301	GACTGCCACAATGGAGGTCAATGTTACCAGCCACAGGGGCGCTGTGAGTGTGAGCCTGGC	360
Qy		1033	TTTGTCTGGCGAGCGCTGCCAAGCACGCCCTGTGTCTGAGGGGCTCTACGGCATCAAATGT	1092
Db		361	TACAAGGGCCCTAGCTGCCAGGAGCGGCTATGCCCTGAGGGCCTGCATGGCCCAGGCTGC	420
Qy		1093	GACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTACCCCATGCTCGGAGAG	1152
Db		421	ACCTTGCCCTGCCCTGTGACACCGAGAACACTATCAGCTGCCATCCAGTTACTGGAGCT	480
Qy		1153	TGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGATTTC	1212
Db		481	TGTACCTGCCAACCAGGCTGGTCTGGCCACTACTGCAATGAGTCCTGCCCCGCCGGCTAC	540
Qy		1213	TACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTG	1272
Db		541	TATGGCAACGGTTGCCAGCTACCCTGCACCTGCCAGAACGGTGCTGACTGCCACAGTATC	600
Qy		1273	ACTGGAAGTGACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCT	1332
Db		601	ACCGGGAGCTGCACTTGTGCTCCAGGCTTCATGGGAGAGGTGTGTGCCGTCCCCTGTGCT	660
Qy		1333	CTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGC	1392
Db		661	GCAGGGACCTATGGTCCCAACTGTTTCATCTGTATGTAGCTGTAGCAACGGCGGCACCTGT	720
Qy		1393	TCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGTGGACTGCTCCATC	1452
Db		721	TCCCCAGTGGATGGCTCCTGCACCTGCCGAGAGGGATGGCAGGGCCTGGACTGCTCCCTG	780
Qy		1453	AGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGG	1512
Db		781	CCTTGTCCCAGTGGGACCTGGGGCCTGAACTGCAATGAGACTTG CATCTGTGCCAATGGA	840
Qy		1513	GGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAA	1572
Db		841	GCTGCCTGCAGCCCCTTTGATGGGTCTGTGCCTGCACCCAGGCTGGCTGGGGGACTCC	900
Qy		1573	TGCGAACTTCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGC	1632
Db		901	TGTGAACTGCCCTGCCCGGACGGCACTTTTGGGCTGAACTGCAGTGAGCATTGCGACTGC	960

Qy 1633 AGCCACGCAGATGGCTGCCACCCTACCACGGGGCCATTGCCGCTGCCTCCCGGGATGGTCA 1692
 |||| | | ||||| |||| | | |||| | | |||| | | |||| | |
 Db 961 AGCCATGCTGATGGCTGTGACCCTGTACAGGCCACTGCTGCTGCCTGGCAGGATGGACA 1020

Qy 1693 GGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCC 1752
 || | | |||| | | || | | ||||| ||||| | | | | | | | | | | | |
 Db 1021 GGCATCCGCTGTGATAGCACGTGTCTCCAGGTCGCTGGGGCCCCAACTGTTTCAGTGTCC 1080

Qy 1753 TGCTACTGTAAAAATGGGGCTTCATGCTCCCTGATGATGGCATCTGCGAGTGTGCACCA 1812
 || | || | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1081 TGCAGCTGTGAGAACGGAGGTTCTGCTCCCCGGAGGACGGGAGCTGCGAGTGTGCCCCCT 1140

Qy 1813 GGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCTGGTTTTTATGGGCATCGCTGC 1872
 |||| | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1141 GGCTTTCGAGGACCCTTATGTGAGAGAATCTGCCCACCAGGATTCACGGCCATGGCTGC 1200

Qy 1873 AGCCAGACATGCCCACAGTGC GTTACAGCAGCGGGCCCTGCCACCACATCACCGGCCTG 1932
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1201 GCCCAGCCTTGTCCCCTCTGCGTGCACAGCAGGGGGCCCTGCCACCACATCAGTGGTATC 1260

Qy 1933 TGTGACTGCTTGCCTGGCTTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCAGTGGCAGA 1992
 |||| | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1261 TGTGAGTGCCTGCCAGGATTCTCTGGAGCCTTGTGCAACCAAGTGTGTGCTGGAGGGCAC 1320

Qy 1993 TTTGGGAAAACTGTGCGAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCAT 2052
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1321 TTCGGGCAGGACTGTGCCAGCTCTGTTCTGTGCCAACAATGGGACCTGCAGCCCCATC 1380

Qy 2053 GACAGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGA CTGCTCTCAACCATGTCCA 2112
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1381 GATGGCTCCTGTGAGTGTCTCCCTGGGTGGATTGGCAAGGACTGCTCACAGGCCTGCCCA 1440

Qy 2113 CCTGCCCACTGGGGCCCAA CTGTCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGC 2172
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1441 TCTGGGTTCTGGGGCTCTGCCTGCTTCCACACATGCAGCTGCCACAACGGGGCGAGCTGC 1500

Qy 2173 AGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAG 2232
 |||| | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1501 AGCGCCGAGGATGGGGCTGCCACTGCACCCCTGGCTGGACTGGACTCTTCTGCACGCAG 1560

Qy 2233 AGATGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACGGA 2292
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1561 CGTTGCCCTTCAGCATTTTTTGGGAAGGACTGTGGGCACATATGCCAGTGTGAGAATGGA 1620

Qy 2293 GCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGCAC 2352
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1621 GCCAGCTGTGACCACATCACTGGGAAATGCACCTGTGGAACAGGCTTCTCTGGCCGCCAC 1680

Qy 2353 TGTGAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGT 2412
 |||| | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1681 TGTGAACAGAGATGTGCCCTTGGAACCTTTGGATATGGGTGTCAGCAGCTATGTGAGTGC 1740

Qy 2413 CTGAACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAG 2472
 |||| | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1741 ATGAACAATGCCACTTGTGACCACGTCACTGGTACCTGTTACTGTAGCCCGGGATTCAAA 1800

Qy 2473 GGAGCGAGATGTGATCAAGCTGGTGTTCATCATAGTTGGAAATCTGAACAGCTTAAGCCGA 2532
 ||| || ||||| ||||| | || | | | |||| | |
 Db 1801 GGAATCAGGTGTGACCAAGCTGCCCTCATGATGGATG---AGCTGAATCCCTACACCAAG 1857

 Qy 2533 ACCAGTACTGCTCTCCCTGCTGATTCTACCAGATCGGGGCCATTGCAGGCATCATCATT 2592
 | |||| | |||| | || | | | | | | ||||| | |
 Db 1858 ATCAGTCCAGCTCTGGGAGCAGAGCGGCACTCAGTGGTGCTGTCACCGGCATCGTTCTC 1917

 Qy 2593 CTTGTCCTAGTTGTTCTCTTCTACTGGCATTGTTTCATTATTTATAGACACAAGCAGAAG 2652
 || | | | || | | || |||| |||| | |||||
 Db 1918 CTGTTGTTCTGGTGGTGGTGCTGCTGGGCCTGTTTGCCTGGCGACGGAGGCGGCAGAAA 1977

 Qy 2653 GGAAAGGAATCAAGCATG---CCAGCAGTTACCTACACCCCTGCTATGAGGGTCGTCAAT 2709
 | || | | || | | || ||||| || ||||| | ||
 Db 1978 GAGAAAGGCCGTGACCTGGCTCCCCGAGTCTCCTACACCCAGCCATGAGGATGACCAGC 2037

 Qy 2710 GCAGATTATACCATTTTCAGGAACCCCTTCTCACAGCAATGGTGAAAC-----GCTAAT 2763
 |||| | | |||| | | |||| | | | | ||||
 Db 2038 ACAGACTACTCTCTCTCAGATTTGTCTCAAAGTAGCAGCCATGCCAGTGCTTTTCCAAT 2097

 Qy 2764 AGCCACTACTTCACCAATCCCAGTTACCACACGCTCACCCAGTGTCACATCCCCCTCAC 2823
 | |||| || | || | | |||| |||| | | ||
 Db 2098 GCCAGCTACCACACACTGGCGTGTGGGGGGCCTGCCACCAGCCAGGCCAGCACTCTGGAC 2157

 Qy 2824 GTCAACAACAGGGACAGGATGACTGTCACG-AAGTCAAAAAACAATCAACTGTTTGTGAA 2882
 |||| | || | | | || | | | | | | | |
 Db 2158 AGGAACAGCCCCACCAAGCTCAGTAACAAGTCCCTTGACAGAGACACAGCAGGCTGGACC 2217

 Qy 2883 TCTTAAAAATGTGAACCCTGGGAAGA-----GAGGCCCTGTGGGGGACTG 2927
 | || | | || | | || | | || || || ||
 Db 2218 CCCTACAGCTATGTGAACGTGTTAGACTCCCATTTCCAGATCAGTGCCCTGGAGGCCAGG 2277

 Qy 2928 CACTGGGACATTGCCGGCTGACTGGAAACATGGCGGCT-----ACCTCAAC 2973
 || | | | | | || | || | | | | | | ||
 Db 2278 TACCCGCCCCGAGGACTTCTACATTGAACTTAGACACCTCAGCCGCCATGCTGAGCCACAC 2337

 Qy 2974 GAGCTCGGTGCTTTTGGACTTGACAGAAG-----CTATATGGGAAAATCCTTA 3021
 | || |||| || | |||| | | |||| |||| ||
 Db 2338 TCACCAGGCATTGTGGAATGGACAGACGTCAGAACACATACATTATGGACAAAGGCTTC 2397

 Qy 3022 AAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAACTGCTCCCTAAGCAGTTCTGAG 3081
 |||| | | |||| | | |||| | | |||| ||
 Db 2398 AAAGATTACATGAAAGAATCTGTGTGCAGTTCTAGCACTTGCTCCTTGAACAGCAGTGAA 2457

 Qy 3082 AACCCATATGCCACTATTAAAGACCCACCTGTACTTATCCCGAAAAGCTCAGAGTGTGGT 3141
 |||| | |||| |||| ||||| | || | | |||| ||
 Db 2458 AACCTTACGCCACAATTAAGGACCCACCCATCCTCACCTGCAAGCTTCCAGAAAGCAGT 2517

 Qy 3142 TATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCATATGCAGAGATCAATAACTCA 3201
 |||| | |||| || || | | | || || || |||| | ||
 Db 2518 TATGTAGAAATGAAGTCACCTGTGCACTTGGGGTCTCCGTACACAGATGTGCCATCCTTG 2577

 Qy 3202 ACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACAGTGAGTGTTGTCCAAGGAGTA 3261
 | || | || | || | ||||| || |||| |||| ||||| ||
 Db 2578 TCGACATCTAATAAAAAATATATATGAAGTTGAGCCACAGTCAGTGTGGTCCAAGAAGGC 2637

 Qy 3262 TTCAGCAATAATGGGCGTCTCTCCCAGGATCCATATGACCTCCCAAAGAACAGTCACATC 3321

```

      |  | | | |  |  | | | | | | | | | | | | | | | | | |
Db      2638 CGAGGTCATAACTCCAGCTATATCCAGAATCCATACGACCTACCTAAGAACAGCCATATT 2697

QY      3322 CCTTGTCATTATGACCTGCTGCCAGTCCGAGACAGTTCATCC 3363
      | | | | | | | | | | | | | | | | | | | | | |
Db      2698 CCTGGTCACTATGACCTCCTCCCAGTAAGACAGAGCCCTGCC 2739

```

RESULT 9

HSM805375

LOCUS HSM805375 3281 bp mRNA linear PRI 12-JUL-2002

DEFINITION Homo sapiens mRNA; cDNA DKFZp434L121 (from clone DKFZp434L121).

ACCESSION AL834326

VERSION AL834326.1 GI:21739945

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3281)

AUTHORS Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and Wiemann,S.

TITLE Direct Submission

JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp434L121) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source

Location/Qualifiers

1..3281

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="RZPD:DKFZp434L121"

/db_xref="taxon:9606"

/clone="DKFZp434L121"

/tissue_type="testis"

/clone_lib="434 (synonym: htes3). Vector pSport1; host

DH10B; sites NotI + SalI"

/dev_stage="adult"

gene

1..3281

/gene="DKFZp434L121"

CDS

<1..1883

/gene="DKFZp434L121"

/note="similarity to MEGF6 (Rattus norvegicus)"

/codon_start=3

/product="hypothetical protein"

/protein_id="CAD38994.1"

/db_xref="GI:21739946"

/db_xref="GOA:Q8ND91"

/db_xref="SPTREMBL:Q8ND91"

/translation="GRVGPGAVCAQPCPPGTFGQNCSDCPCHGGQCDHVTGQCHCT
AGYMGDRCQEECPFGSFGFQCSQRCDCHNGGQCSPTTGACECEPGYKGPRCQERLCPE

Qy	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTA	1440
Db	618	GGTGGCACCTGCTCCCCAGTAGATGGCTCCTGTAC	677
Qy	1441	GACTGCTCCATCAGATGTCCCAGTGGCACATGGGG	1500
Db	678	GACTGCACCCTGCCATGTCCCAGTGGGACGTGGGG	737
Qy	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGG	1560
Db	738	TGTGCCAATGGGGCAGCCTGCAGCCCCATAGACGG	797
Qy	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGG	1620
Db	798	CTGGGAGACACCTGTGAGCTGCCTTGCCCGGATGG	857
Qy	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACC	1680
Db	858	CACTGTGACTGCAGCCATGCTGATGGATGTGACCC	917
Qy	1681	CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGT	1740
Db	918	GCCGGATGGACAGGCATCCGCTGTGACAGCACGT	977
Qy	1741	TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTT	1800
Db	978	TGCTCTGTCTCCTGCAGCTGTGAGAATGGAGGCT	1037
Qy	1801	GAGTGTGCACCAGGCTTCCGAGGCACCACTGTG	1860
Db	1038	GAGTGTGCCCTGGCTTCCGAGGACCCTTATGCC	1097
Qy	1861	GGGCATCGCTGCAGCCAGACATGCCACAGTGCGT	1920
Db	1098	GGCCATGGCTGCGCCCAGCCATGCCCCCTCTGCG	1157
Qy	1921	ATCACCGGCCTGTGTGACTGCTTGCCTGGCTTCA	1980
Db	1158	ATCAGCGGCATCTGTGAGTGCCTCCAGGATTCT	1217
Qy	1981	CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAAT	2040
Db	1218	GCTGGAGGATACTTTGGGCAGGACTGTGCCAGCT	1277
Qy	2041	TGTAACCCCATTGACAGATCTTGTCA	2100
Db	1278	TGCAGCCCTATCGATGGCTCCTGCCAGTGCTTTC	1337
Qy	2101	CAACCATGTCCACCTGCCCACTGGGGCCAACTGC	2160
Db	1338	CAGGCTTGCCACCCGGGTTC	1397
Qy	2161	GGAGCTTCTGCGAGCGCCTACGATGGGGAA	2220
Db	1398	GGGGCGAGCTGCAGCGCCGAGGACGGGGCCTGCC	1457

Qy 2221 TACTGCACTCAGAGATGTCCTC 2242
| | | | | | | | | | | | | |
Db 1458 TTCTGCACACAGCGTAAGCCCC 1479

RESULT 10

AF444274

LOCUS AF444274 4290 bp mRNA linear ROD 06-DEC-2001

DEFINITION Mus musculus Jedi protein mRNA, complete cds.

ACCESSION AF444274

VERSION AF444274.1 GI:17386052

KEYWORDS .

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 4290)

AUTHORS Krivtsov,A.V., Zinovyeva,M.V., Hendrikx,J., Visser,J.W.M. and
Belyavsky,A.V.

TITLE Jedi is a novel DSL and EGF-like repeat motif-containing protein
expressed on non-differentiated hematopoietic cells

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 4290)

AUTHORS Krivtsov,A.V., Zinovyeva,M.V., Visser,J.W.M. and Belyavsky,A.V.

TITLE Direct Submission

JOURNAL Submitted (07-NOV-2001) Stem Cell Biology, Lindsley F. Kimball
Research Institute, New York Blood Center, 310 East 67 Street, New
York, NY 10021, USA

FEATURES Location/Qualifiers

source 1. .4290

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL"

/db_xref="taxon:10090"

/clone="E-7"

/tissue_type="testis"

CDS 138. .3242

/codon_start=1

/product="Jedi protein"

/protein_id="AAL38571.1"

/db_xref="GI:17386053"

/translation="MPLCPLLLLALGLRLTGTLNSNDPNVCTFWESFTTTTKESHLRP
FSLLPAESCHRPWEDPHTCAQPTVVYRTVYRQVVKMDSRPRLQCCRGYYESRGACVPL
CAQECVHGRCVAPNQCCAPGWRGGDCSSECAPGMWGPQCDKFCHCGNNSCDPKSGA
CFCPSGLQPPNCLQPCPAGHYGPACQFDCQCYGASCDPQDGACFCPPGRAGPSCNVPC
SQGTDGFFCPRTYPCQNGGVPQGSQGSQCSPPGWMGVICSLPCPEGFHGPNCTQECRC
HNGGLCDRFTGQCHCAPGYIGDRQEECPVGRFGQDCAETCDCAPGARCFPANGACLC
EHGFTGDRCTERLCPDGRYGLSCQEPCTCDPEHSLSCHPMHGECSCQPGWAGLHCNES
CPQDTHGPGCQEHCLCLHGGGLCLADSGLCRCAPGYTGPHCANLCPDITYGINCSSRCS
CENAIACSPIDGTCICKEGWQRGNCSVPCPLGTWGFNCNASCQCAHDGVCSPQTGACT
CTPGWHGAHCQLPCPKGQFGEGCASVDCDHS DGCDPVHGQCRCQAGWMGTRCHLPCP
EGFWGANCSNTCTCKNGGTCVSENGNCVCA PGFRGPSCQRPCPPGRYGRKCVQCKCNN
NHSSCHPSDGTCSCLAGWTGPDCEACPPGHWGLKCSQLCQCHHGGTCHPQDGSCICT
PGWTGPNCLEGCPRMFGVNCSQLCQCDLGEMCHPQTGACVCPPGHSGADCKMGSQES
FTIMPTSPVTHNSLGAVIGIAVLGTLVVALIALFIGYRQWQKGKEHEHLAVAYSTGRL
DGS DYVMPDVSPSYSHYYSNPSYHTLSQCSPNPPPNKVPGSQLFVSSQAPERPSRAH
GRENHVTLPADWKHRRPHERGASHLDRSYSCSYSHRNGPGPFCHKGPIS E EGLGASV

Qy 671 TTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAAATGGAG 730
 ||| | | || | |||| | | | |||| ||||| |||||
 Db 787 CCTGTTACAGGGCACTGATGGCTTCTTCTGCCCCAGAACCTATCCTTGCCAAATGGAG 846

Qy 731 GAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGT 790
 | || ||| | | || ||| |||| | ||||| ||||| | |
 Db 847 GTGTTCCCTCAGGGCTCCCAAGGCTCCTGCAGCTGCCACCGGGCTGGATGGGTGTCATT 906

Qy 791 GTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCC 850
 || | ||| |||| ||||| | |||| ||||| | | |||| | ||||
 Db 907 GTTCCCTGCCATGCCAGAGGGTTTCCATGGACCCAAGTGTACTCAGGAATGTCGCTGCC 966

Qy 851 ATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAG 910
 | || | || |||| | || | || | || | || | || | || |
 Db 967 ACAACGGTGGCCTCTGTGACAGGTTTACTGGGCAGTGCCACTGTGCTCCTGGCTATATCG 1026

Qy 911 GGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCT 970
 |||| ||||| || ||||| || || || || || || ||||| |||||
 Db 1027 GGGATCGGTGCCAAGAAGAGTGCCCCGTGGGCCGCTTTGGTCAAGACTGTGCTGAGACCT 1086

Qy 971 GCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAG 1030
 | | |||| || | || | | | | |||| | || | || |
 Db 1087 GTGACTGTGCTCCTGGCGCCCGTTGCTTTCCTGCTAATGGCGCGTGTCTGTGCGAACATG 1146

Qy 1031 GCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTCTGAGGGGCTCTACGGCATCAAAT 1090
 |||| | |||| ||||| | |||| |||| | || | || | || |
 Db 1147 GCTTCACAGGCGACCGCTGCACTGAGCGCCTCTGTCCGGATGGCCGCTATGGTCTGAGCT 1206

Qy 1091 GTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTCACCCCATGTCTGGAG 1150
 | | | | || |||| || || || || || |||| |||| | || |
 Db 1207 GCCAGGAGCCCTGCACCTGCGACCCAGAACACAGTCTCAGCTGCCACCCGATGCACGGCG 1266

Qy 1151 AGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGAT 1210
 |||| |||| | |||| || || | || || |||| | |||| || || |
 Db 1267 AGTGCTCCTGCCAGCCAGGTTGGGCGGGCCTCCACTGCAACGAGAGCTGCCCTCAGGACA 1326

Qy 1211 TCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAATGGGGCAGACTGTGACAGTG 1270
 | || | ||||| || |||| || | | || | || | || |
 Db 1327 CGCATGGCCCGGGCTGCCAGGAGCACTGCCTCTGTCTGCACGGAGGGCTCTGCCTTGCCG 1386

Qy 1271 TGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCC 1330
 | || || || || || ||||| || || | |||| || || || || |
 Db 1387 ACAGCGGCCTCTGTGCGTGCGGCCAGGATACAGGGACCTCACTGCGCTAACCTATGTC 1446

Qy 1331 CTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCT 1390
 | | || || ||||| ||||| |||| | |||| |||| | || |
 Db 1447 CACCGGACACTTATGGGATCAACTGTTCCCTCCCGCTGCTCCTGTGAAATGCCATTGCCT 1506

Qy 1391 GCTCTCCTGTGGACGGGTCTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCA 1450
 ||||| | |||| | || | ||||| || |||| | | |||||
 Db 1507 GCTCTCCCATTTACGGCACGTGCATCTGCAAGGAAGTTGGCAGCGTGGTAACTGCTCTG 1566

Qy 1451 TCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCCTCAACG 1510
 | ||||| ||||| ||||| || || | ||||| | || |
 Db 1567 TTCCCTGTCCCTTGGCACCTGGGGCTTCAATTGCAATGCCAGTTGCCAGTGTGCCACG 1626

Qy 1511 GGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGGGGGAGA 1570

```

      ||| |||| | |||   || |||| || || | |||| |||| || |
Db      1627 ATGGAGTCTGCAGCCCCAACTGGAGCCTGTACTTGCACCCCTGGGTGGCATGGTGCTC 1686
Qy      1571 AATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACT 1630
      | ||| | ||||| ||| | || || || ||||| ||| ||||
Db      1687 ACTGCCAGCTTCCCTGCCCGAAGGGACAGTTTGGTGAAGGCTGTGCCAGTGTCTGTGACT 1746
Qy      1631 GCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGT 1690
      | |||| | ||||| |||| | || |||| || | | || |||
Db      1747 GTGACCACTCTGATGGCTGTGACCCTGTTTCATGGACAGTGCCGATGTCAGGCTGGTTGGA 1806
Qy      1691 CAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGC 1750
      || | ||| || || | |||| | |||| | |||||
Db      1807 TGGGCACACGCTGCCACCTGCCTTGCCCGAGGGCTTTTGGGGAGCCAACCTGCAGTAACA 1866
Qy      1751 CCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCAC 1810
      |||| ||| || |||| | | || || |||| | |||| | ||||
Db      1867 CCTGTACCTGCAAGAATGGTGGTACCTGTGTGTCTGAGAATGGCAACTGCGTGTGCGCAC 1926
Qy      1811 CAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCT 1870
      |||| | ||||| || | || |||| | |||| | |||| | ||||
Db      1927 CAGGGTTCCGAGGCCCTCCTGCCAGAGGCCCTGCCCCGCTGGTCGCTATGGCAAACGCT 1986
Qy      1871 GCAGCCAGACATGCCACAGTGCCTTCACAGCAGCGGGGCCCTGCCACCACATCACCGGCC 1930
      | || || |||| ||| | ||||| ||||
Db      1987 GTGTGCA-----ATGCAAGTGTAACAACAACCATTCTTCCTGCCACCCATCGGACGGGA 2040
Qy      1931 TGTGTGACTGCTTGCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCA 1990
      || |||| || | |||| |||| | |||| ||| | |||| |||
Db      2041 CCTGCTCCTGCCTGGCGGGCTGGACAGGCCCTGACTGCTCTGAGGCATGTCCCCCAGGCC 2100
Qy      1991 GATTTGGGAAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCA 2050
      | || || || | || || || || || || || || || || || ||
Db      2101 ACTGGGGACTCAAATGCTCCCAACTCTGCCAGTGTGTCATCATGGTGGGACCTGCCACCCCC 2160
Qy      2051 TTGACAGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTCTCAACCATGTC 2110
      || | ||| || || || || |||| ||| ||||| || |||
Db      2161 AGGATGGGAGCTGTATCTGCACGCCAGGCTGGACTGGACCAACTGCTTGAAGGCTGCC 2220
Qy      2111 CACCTGCCCACTGGGGCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCT 2170
      |||| | || ||||| ||| || || || || || || || ||
Db      2221 CACCAAGAATGTTTGGTGTCAACTGCTCCCAGCTATGTGAGTGTGATCTCGGAGAGATGT 2280
Qy      2171 GCAGCGCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTC 2230
      || || | |||| ||| || || || || | || |||||
Db      2281 GCCACCCACAGACTGGGGCTTGTGTCTGTCCCCCAGGACACAGTGGTGCAGACTGCAAAA 2340
Qy      2231 AGAGATGTC 2239
      | ||| |
Db      2341 TGGGAAGCC 2349

```

RESULT 11

AF461685

LOCUS AF461685 4539 bp mRNA linear ROD 21-JAN-2002

DEFINITION Mus musculus Jedi-736 protein mRNA, complete cds.

ACCESSION AF461685

Db	183	AGGAGTCCCACCTGCGCCCCTTCAGCCTGCTCCCAGCTGAGTCCTGCCACAGGCCCTGGG	242
Qy	191	TAAACTGGTTTAAATGCACGGGCACAGAGTCAGCTATCGGACAGCCTATCGACATGGGG	250
Db	243	AGGACCCCCACACCTGTGCCCAGCCTACGGTTGTCTACCGGACTGTGTACCGTCAGGTGG	302
Qy	251	AGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCTGGATTTTATGAAAGCGGGGAAA	310
Db	303	TGAAGATGGACTCCCCGCCACGCCTGCAGTGCTGTAGGGGTACTACGAGAGCAGAGGGG	362
Qy	311	TGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCTCCAAACACCT	370
Db	363	CCTGTGTCCCACTCTGTGCCCAGGAGTGTGTCCATGGTCGCTGTGTGGCTCCGAATCAGT	422
Qy	371	GTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCGTGCGATGGTGATCACT	430
Db	423	GCCAGTGTGCACCAGGCTGGCGGGGTGGCGACTGCTCCAGCGAGTGTGCCCCGGAATGT	482
Qy	431	GGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAATGGGGCTCTGTGCAACCCCATCA	490
Db	483	GGGGACCACAGTGTGACAAGTTCTGCCACTGTGGCAACAACAGTTCCTGTGATCCCAAGA	542
Qy	491	CCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGC	550
Db	543	GTGGGGCTGCTTTTGCCCCCTCTGGCCTGCAGCCCCCAACTGCCTTCAGCCTTGCCCTG	602
Qy	551	AGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCG	610
Db	603	CCGGCCACTATGGTCCTGCCTGCCAGTTTGATTGCCAGTGC---TATGGGGCATCCTGTG	659
Qy	611	ACCACGTCACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTCTGTGAGGATC	670
Db	660	ACCCCCAGGATGGAGCCTGTTTCTGCCCTCCAGGGAGAGCAGGACCCAGCTGTAATGTGC	719
Qy	671	TTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAAATGGAG	730
Db	720	CCTGTTACAGGGCACTGATGGCTTCTTCTGCCCCAGAACCTATCCTTGCCAAAATGGAG	779
Qy	731	GAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGT	790
Db	780	GTGTTCTCAGGGCTCCCAAGGCTCCTGCAGCTGCCCACCGGGCTGGATGGGTGTCATT	839
Qy	791	GTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCC	850
Db	840	GTTCCCTGCCATGCCAGAGGGTTTCCATGGACCCAAGTGTACTCAGGAATGTCGCTGCC	899
Qy	851	ATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAG	910
Db	900	ACAACGGTGGCCTCTGTGACAGGTTTACTGGGCAGTGCCACTGTGCTCCTGGCTATATCG	959
Qy	911	GGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCT	970
Db	960	GGGATCGGTGCCAAGAAGAGTGCCCCGTGGGCCGCTTTGGTCAAGACTGTGCTGAGACCT	1019
Qy	971	GCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAG	1030

Db 1020 GTGACTGTGCTCCTGGCGCCCGTTGCTTTCCTGCTAATGGCGCGTGTCTGTGCGAACATG 1079

Qy 1031 GCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGCTCTACGGCATCAAAT 1090
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1080 GCTTCACAGGCGACCGCTGCACTGAGCGCCTCTGTCCGGATGGCCGCTATGGTCTGAGCT 1139

Qy 1091 GTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTCACCCCATGTCTGGAG 1150
 | | | | | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1140 GCCAGGAGCCCTGCACCTGCGACCCAGAACACAGTCTCAGCTGCCACCCGATGCACGGCG 1199

Qy 1151 AGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGAT 1210
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1200 AGTGCTCCTGCCAGCCAGGTTGGGCGGGCCTCCACTGCAACGAGAGCTGCCCTCAGGACA 1259

Qy 1211 TCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTG 1270
 | || | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1260 CGCATGGCCCCGGGCTGCCAGGAGCACTGCCTCTGTCTGCACGGAGGGCTCTGCCTTGCCG 1319

Qy 1271 TGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCC 1330
 | || | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1320 ACAGCGGCCTCTGTGCGGTGCGCGCCAGGATACAGGGACCTCACTGCGCTAACCTATGTC 1379

Qy 1331 CTCTGGGAACCTATGGGATAAACTGTTTCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCT 1390
 | | || | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1380 CACCGGACACTTATGGGATCAACTGTTTCCTCCGCTGCTCCTGTGAAAATGCCATTGCCCT 1439

Qy 1391 GCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCA 1450
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1440 GCTCTCCCATTGACGGCACGTGCATCTGCAAGGAAGGTTGGCAGCGTGGTAAC'TGCTCTG 1499

Qy 1451 TCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAAC'TAACATGCCAGTGCCTCAACG 1510
 | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1500 TTCCCTGTCCCTTGGCACCTGGGGCTTCAATTGCAATGCCAGTTGCCAGTGTGCCACAG 1559

Qy 1511 GGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGA 1570
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1560 ATGGAGTCTGCAGCCCCCAAACCTGGAGCCTGTACTTGCACCCCTGGGTGGCATGGTGCTC 1619

Qy 1571 AATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAAC'TGTGCTGAGCGCTGCGACT 1630
 | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1620 ACTGCCAGCTTCCCTGCCCCAAGGGACAGTTTGGTGAAGGCTGTGCCAGTGTCTGTGACT 1679

Qy 1631 GCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGT 1690
 | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1680 GTGACCACTCTGATGGCTGTGACCCTGTTTCATGGACAGTGCCGATGTCAGGCTGGTTGA 1739

Qy 1691 CAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGC 1750
 || | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1740 TGGGCACACGCTGCCACCTGCCTTGCCCGAGGGCTTTTGGGGAGCCAAC'TGCAGTAACA 1799

Qy 1751 CCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCAC 1810
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1800 CCTGTACCTGCAAGAATGGTGGTACCTGTGTGTCTGAGAATGGCAACTGCGTGTGCGCAC 1859

Qy 1811 CAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTATGGGCATCGCT 1870
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 1860 CAGGGTTCGAGGCCCTCCTGCCAGAGGCCCTGCCCGCTGGTCGCTATGGCAAACGCT 1919

Qy 1871 GCAGCCAGACATGCCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCC 1930
 | || ||| ||| | ||| |||
 Db 1920 GTGTGCA-----ATGCAAGTGTAAACAACAACCATTCTTCCTGCCACCCATCGGACGGGA 1973

Qy 1931 TGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCAGTGGCA 1990
 || ||| || | ||| ||| | ||| ||| | ||| |||
 Db 1974 CCTGCTCCTGCCTGGCGGGCTGGACAGGCCCTGACTGCTCTGAGGCATGTCCCCAGGCC 2033

Qy 1991 GATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCA 2050
 | || || || | | || || | || ||| |||
 Db 2034 ACTGGGACTCAAATGCTCCCAACTCTGCCAGTGTGCATCATGGTGGGACCTGCCACCCCC 2093

Qy 2051 TTGACAGATCTTGTCACTGTTACCCCGGTTGGATTGGCAGTGAAGTGTCTCAACCATGTC 2110
 || | ||| || || || ||| ||| ||| || ||
 Db 2094 AGGATGGGAGCTGTATCTGCACGCCAGGCTGGACTGGACCCAACTGCTTGAAGGCTGCC 2153

Qy 2111 CACCTGCCCCTGCGGGCCCAAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCT 2170
 ||| | || ||| ||| || || || || ||
 Db 2154 CACCAAGAATGTTTGGTGTCAACTGCTCCCAGCTATGTCACTGTGATCTCGGAGAGATGT 2213

Qy 2171 GCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCA 2227
 || || | ||| || || || || || ||
 Db 2214 GCCACCCACAGACTGGGGCTTGTGTCTGTCCCCCAGGACACAGTGGTGCAGACTGCA 2270

RESULT 12

AF440279

LOCUS AF440279 4482 bp mRNA linear ROD 20-NOV-2001

DEFINITION Mus musculus MEGF12 (Megf12) mRNA, complete cds.

ACCESSION AF440279

VERSION AF440279.1 GI:17017250

KEYWORDS .

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 4482)

AUTHORS Ivanova,N.B. and Lemischka,I.R.

TITLE The global gene expression profiling of the hematopoietic stem cell

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 4482)

AUTHORS Ivanova,N.B. and Lemischka,I.R.

TITLE Direct Submission

JOURNAL Submitted (25-OCT-2001) Molecular Biology, Princeton University,
 Washington Road, Princeton, NJ 08544, USA

FEATURES Location/Qualifiers

source

1. .4482

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/cell_type="AA4+Sca+Kit+Lin- hematopoietic stem cell"

/tissue_type="liver"

/dev_stage="14 day old fetus"

gene

1. .4482

/gene="Megf12"

CDS

338. .3442
 /gene="Megf12"
 /note="contains signal peptide, 13.5 EGF-like-domains and transmembrane domain"
 /codon_start=1
 /product="MEGF12"
 /protein_id="AAL33583.1"
 /db_xref="GI:17017251"
 /translation="MPLCPLLLLALGLRLTGTLNSNDPNVCTFWESFTTTTKESHLRP
 FSLLPAESCHRPWEDPHTCAQPTVVYRTVYRQVVKMDSRPRLQCCRGYYESRGACVPL
 CAQECVHGRCVAPNQCCAPGWRGGDCSSECAPGMWGPQCDKFCHCGNNSCDEPKSGT
 CFCPSGLQPPNCLQPCPAGHYGPACQFDCQCYGASCDPQDGACFCPPGRAGPSCNVPC
 SQGTDGFFCPRTYPCQNGGVPQGSQGSQSCSPPGWMGVICSLPCPEGFHGPNCTQECRC
 HNGGLCDRFTGQCHCAPGYIGDRCQEECPVGRFGQDCAETCDCAPGARCFFPANGACLC
 EHGF TGDRCTERLCPDGRYGLSCQEPCTCDPEHSLSCHPMHGECSCQPGWAGLHCNES
 CPQDTHGPGCQEHCLCLHGGCLCLADSGLCRCAPGYTGPHCANLCPDPTYGINCSSRCS
 CENAIACSPIDGTICKEGWQRGNCSVPCPLGTWGFNCNASCQCAHDGVCSFQTGACT
 CTPGWHGAHCQLPCPKGQFGEGCASVDCDHDSDGCDPVHGGQCRCAQAGWMGTRCHLPCP
 EGFWGANCSNTCTCKNGGTCVSENGNCVCAPGFRGPSCQRPCCPGRYGKRCVQCKCNN
 NHSSCHPSDGTCSCLAGWTGPDCEACPPGHWGLKCSQLCQCHHGGTCHPQDGSCICT
 PGWTGPNCLEGCPPRMFGVNCSQLCQCDLGEMCHPETGACVCPPGHSGADCKMGSQES
 FTIMPTSPVTHNSLGA VIGIAVLGTLVVALIALFIGYRQWQKGKEHEHLAVAYSTGRL
 DGSDYVMPDVSPSYSHYYSNPSYHTLSQCSNPPPPNKVPGSQLFVSSQAPERPSRAH
 GRENHVTLPADWKHRRPHERGASHLDRSYSCSYSHRNGPGPFCHKGPISSEGLGASV
 MSLSENPYATIRDLP SLPGEPRESGYVEMKGPPSVSPRQSLHLRDRQQRQLQPQRD
 SGTYEQPSPLSHNEESLGSTPPLPPGLPPGHYDSPKNSHIPGHYDLFPVRHPPSPPSR
 RQDR"

ORIGIN

Query Match 18.9%; Score 646.2; DB 10; Length 4482;
 Best Local Similarity 57.1%; Pred. No. 5e-188;
 Matches 1239; Conservative 0; Mismatches 918; Indels 12; Gaps 3;

Qy	74	CTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACTCAGTGACTGTGC	133
Db	390	CACTCAACTCCAATGATCCCAATGTCTGTACCTTTTGGGAAAGCTTCACCACGACCACTA	449
Qy	134	AAGAGTCATACCCACATCCCTTTGATCAAAT---TTACTACACGAGCTGCACTGACATTC	190
Db	450	AGGAGTCCCACCTGCGCCCCCTTCAGCCTGCTCCAGCTGAGTCCTGCCACAGGCCCTGGG	509
Qy	191	TAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTATCGACATGGGG	250
Db	510	AGGACCCCCACACCTGTGCCAGCCTACGGTTGTCTACCGGACTGTGTACCGTCAGGTGG	569
Qy	251	AGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCTCGGATTTTATGAAAGCGGGGAAA	310
Db	570	TGAAGATGGACTCCCGCCCACGCCTGCAGTGCTGTAGGGTTACTACGAGAGCAGAGGGG	629
Qy	311	TGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCTCCAAACACCT	370
Db	630	CCTGTGTCCCCACTCTGTGCCAGGAGTGTGTCCATGGTCGCTGTGTGGCTCCGAATCAGT	689
Qy	371	GTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGATGGTGATCACT	430
Db	690	GCCAGTGTGCACCAGGCTGGCGGGGTGGCGACTGCTCCAGCGAGTGTGCCCCGGAATGT	749

Qy	431	GGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGCAACCCCATCA	490
Db	750	GGGGACCACAGTGTGACAAGTTCTGCCACTGTGGCAACAACAGTTCCTGTGATCCCAAGA	809
Qy	491	CCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGC	550
Db	810	GTGGGACGTGCTTTTGCCCTCTGGCCTGCAGCCCCCAACTGCCTTCAGCCCTGCCCTG	869
Qy	551	AGGGCACCTATGGTAACGACTGTTCATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCG	610
Db	870	CCGGCCACTATGGTCCTGCCTGCCAGTTTGATTGCCAGTGC---TATGGGGCATCCTGTG	926
Qy	611	ACCACGTACAGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTCTGTGAGGATC	670
Db	927	ACCCCCAGGATGGAGCCTGTTTCTGCCCTCCAGGGAGAGCAGGACCCAGCTGTAATGTGC	986
Qy	671	TTTGTCTCTCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAAATGGAG	730
Db	987	CCTGTTCACAGGGCACTGATGGCTTCTTCTGCCCCAGAACCTATCCTTGCCAAAATGGAG	1046
Qy	731	GAGTGTGTTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGT	790
Db	1047	GTGTTCCCTCAGGGCTCCCAAGGCTCCTGCAGCTGCCACCAGGGCTGGATGGGTGTCATTT	1106
Qy	791	GTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCC	850
Db	1107	GTTCCCTGCCATGCCAGAGGGTTTCCATGGACCAACTGTACTCAGGAATGTCGCTGCC	1166
Qy	851	ATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAG	910
Db	1167	ACAACGGTGGCCTCTGTGACAGGTTTACTGGGCAGTGCCACTGTGCTCCTGGCTATATCG	1226
Qy	911	GGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCT	970
Db	1227	GGGATCGGTGCCAAGAAGAGTGCCCCGTGGGCCGCTTCGGTCAAGACTGTGCTGAGACCT	1286
Qy	971	GCCAGTGTGTCAACGGAGGGAAGTGTTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAG	1030
Db	1287	GTGACTGTGCTCCTGGCGCCCGTTGCTTTCTGCTAATGGCGCGTGTCTGTGTGAACATG	1346
Qy	1031	GCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTACGGCATCAAAT	1090
Db	1347	GCTTCACAGGCGACCGCTGCACTGAGCGCCTCTGTCCGGATGGCCGCTATGGTCTGAGCT	1406
Qy	1091	GTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTACCCCCATGTCTGGAG	1150
Db	1407	GCCAGGAGCCCTGCACCTGCGACCCAGAACACAGTCTCAGTGCCACCCGATGCACGGCG	1466
Qy	1151	AGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGAT	1210
Db	1467	AGTGCTCCTGCCAGCCAGGTTGGGCGGGCCTCCACTGCAACGAGAGCTGCCCTCAGGACA	1526
Qy	1211	TCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTG	1270
Db	1527	CGCATGGCCCCGGGCTGCCAGGAGCACTGCCTCTGTCTGCACGGAGGGCTCTGCCTTGCCG	1586
Qy	1271	TGACTGGAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCC	1330

Db	1587	ACAGCGGCCTCTGCCGGTGC	CGCGCCGGGATACACGGGACCTCACTGCGCTAACCTATGTC	1646
Qy	1331	CTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAATGATGCAGTCT	1390	
Db	1647	CACCGGACACTTACGGGATCAACTGTTCTCCCGCTGCTCCTGTGAAAATGCCATTGCCT	1706	
Qy	1391	GCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCA	1450	
Db	1707	GCTCTCCCATCGACGGCACGTGCATCTGCAAGGAAGGTTGGCAGCGTGGTAAGTGTCTG	1766	
Qy	1451	TCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACTTAACATGCCAGTGCCTCAACG	1510	
Db	1767	TTCCCTGTCCCCTTGGCACCTGGGGCTTCAATTGCAATGCCAGTTGCCAGTGTGCCACG	1826	
Qy	1511	GGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGA	1570	
Db	1827	ACGGAGTCTGCAGCCCCAACTGGAGCCTGTACTTGACCCCTGGGTGGCATGGTGCTC	1886	
Qy	1571	AATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACT	1630	
Db	1887	ACTGCCAGCTTCCCTGCCCGAAGGGACAGTTTGGTGAAGGCTGTGCCAGTGTCTGTGACT	1946	
Qy	1631	GCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGT	1690	
Db	1947	GTGACCACTCTGATGGCTGTGACCCTGTTTCATGGACAGTGCCGATGTCAGGCTGGTTGGA	2006	
Qy	1691	CAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGC	1750	
Db	2007	TGGGCACACGTGCCACCTGCCTTGCCCGGAGGGCTTTTGGGGAGCCAAGTGCAGTAACA	2066	
Qy	1751	CCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCAC	1810	
Db	2067	CCTGTACCTGCAAGAATGGTGGTACCTGTGTGTCTGAGAATGGCAACTGCGTGTGCGCAC	2126	
Qy	1811	CAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCT	1870	
Db	2127	CAGGGTTCCGAGGCCCTCCTGCCAGAGGCCCTGCCCGCCTGGTCGCTATGGCAAACGCT	2186	
Qy	1871	GCAGCCAGACATGCCACAGTGCGTTTACAGCAGCGGGCCCTGCCACCACATACCCGGCC	1930	
Db	2187	GTGTGCA-----ATGCAAGTGTAACAACAACCATTCCTCCTGCCACCCATCGGACGGGA	2240	
Qy	1931	TGTGTGACTGCTTGCCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCAGTGGCA	1990	
Db	2241	CCTGCTCCTGCCTGGCGGGCTGGACAGGCCCTGACTGCTCCGAGGCATGTCCCCAGGCC	2300	
Qy	1991	GATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCA	2050	
Db	2301	ACTGGGGACTCAAATGCTCCCAACTCTGCCAGTGTATCATGGTGGGACCTGCCACCCCC	2360	
Qy	2051	TTGACAGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTCTCAACCATGTC	2110	
Db	2361	AGGATGGGAGCTGTATCTGCACGCCAGGCTGGACTGGACCAACTGCTTGAAGGCTGCC	2420	
Qy	2111	CACCTGCCCACTGGGGCCCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCT	2170	

Qy	62	GGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACT	121
Db	289	GGCTGGCTGGAAC TCTCAACCCCAGTGATCCAATACCTGCAGCTTCT'GGGAAAGCTTCA	348
Qy	122	CAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC-	180
Db	349	CTACCACCACCAAGGAGTCCCCTCCCGCCCCCTTCAGCCTGCTCCCCTCAGAGCCCTGCG	408
Qy	181	--ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCT	238
Db	409	AGCGGCCCTGGGAGGGCCCCCATACTTGCCCCCAGCCACGGTTGTATACCGGACCGTGT	468
Qy	239	ATCGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTGTCTCCTGGATTTTATG	298
Db	469	ACCGTCAGGTGGTGAAGACGGACCACCGCCAGCGCCTGCAGTGTGCCATGGCTTCTATG	528

Qy	299	AAAGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAAATGTGTCCATGGTCGCTGTATTG	358
Db	529	AGAGCAGGGGGTTCTGTGTCCCGCTCTGTGCCCAGGAGTGTGTCCATGGCCGTTGTGTGG	588
Qy	359	CTCCAAACACCTGTCTAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCG	418
Db	589	CACCCAATCAGTGCCAATGTGTGCCAGGCTGGCGGGGCGACGACTGTTCCAGTGAGTGTG	648
Qy	419	ATGGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGT	478
Db	649	CCCCAGGAATGTGGGGGCCACAGTGTGACAAGCCCTGCAGCTGCGGCAACAACAGCTCGT	708
Qy	479	GCAACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGG	538
Db	709	GTGATCCCAAGAGTGGGGTATGTTCTTGCCCTTCTGGTCTGCAGCCCCGAAGTGCCTTC	768
Qy	539	ACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATG	598
Db	769	AGCCCTGTACCCCTGGCTACTATGGCCCTGCCTGCCAGTTCCGCTGCCAGTGCC---ATG	825
Qy	599	GAGCCACCTGCGACCACGTACAGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCT	658
Db	826	GGGACCCCTGCGATCCCAGACTGGAGCCTGCTTCTGCCCCGAGAGAGAACTGGGCCCA	885
Qy	659	TCTGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTT	718
Db	886	GCTGTGACGTGTCTCTGTTCCAGGGCACTTCTGGCTTCTTCTGCCCCAGCACCCATCCTT	945
Qy	719	GTCAAAATGGAGGAGTGTGTTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGA	778
Db	946	GCCAAAATGGAGGTGTCTTCCAAACCCACAGGGCTCCTGCAGCTGCCCCCTGGCTGGA	1005
Qy	779	TGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAG	838
Db	1006	TGGGCACCATCTGCTCCCTGCCCTGCCAGAGGGCTTTCACGGACCCAACTGCTCCCAGG	1065
Qy	839	AATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTC	898
Db	1066	AATGTCGCTGCCACAACGGCGGCCCTCTGTGACCGATTCACTGGGCAGTGCCGCTGCGCTC	1125
Qy	899	CAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCT	958
Db	1126	CGGGTTACACTGGGGATCGGTGCCGGGAGGAGTGCCCGGTGGGCGCTTTGGGCAGGACT	1185
Qy	959	GTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCC	1018
Db	1186	GTGCTGAGACGTGCGACTGCGCCCCGACGCCCGTTGCTTCCCGCCAACGGCGCATGTC	1245
Qy	1019	TCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCTGTGTCTTGAGGGGCTCT	1078
Db	1246	TGTGCGAACACGGCTTCACTGGGGACCGCTGCACGGATCGCCTCTGCCCCGACGGCTTCT	1305
Qy	1079	ACGGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTCACC	1138
Db	1306	ACGGTCTCAGCTGCCAGGCCCCCTGCACCTGCGACCGGGAGCACAGCCTCAGCTGCCACC	1365

Qy 1139 CCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACAT 1198
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1366 CGATGAACGGGGAGTGTCTCTGCCTGCCGGGCTGGGCGGGCCTCCACTGCAACGAGAGCT 1425

Qy 1199 GTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAG 1258
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1426 GCCCGCAGGACACGCATGGGCCAGGCTGCCAGGAGCACTGTCTCTGCCTGCACGGTGGCG 1485

Qy 1259 ACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCT 1318
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1486 TCTGCCAGGCTACCAGCGGCTCTGTCACTGCGCGCCGGGTACACGGGCCCTCACTGTG 1545

Qy 1319 CTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAA 1378
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1546 CTAGTCTTTGTCTCTGACACCTACGGTGTCAACTGTTCTGCACGCTGCTCATGTGAAA 1605

Qy 1379 ATGATGCAGTCTGCTCTCCTGTGGACGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGG 1438
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1606 ATGCCATCGCCTGCTCACCCATCGACGGCGAGTGCCTGCAAGGAAGGTTGGCAGCGTG 1665

Qy 1439 TGGACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCC 1498
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1666 GTAAGTCTCTGTGCCCTGCCACCCGGAACCTGGGGCTTCAGTTGCAATGCCAGCTGCC 1725

Qy 1499 AGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGAT 1558
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1726 AGTGTGCCCATGAGGCAGTCTGCAGCCCCAACTGGAGCCTGTACCTGCACCCCTGGGT 1785

Qy 1559 GGC CGGGGAGAAATGCGAATTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTG 1618
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1786 GGCATGGGGCCCACTGCCAGCTGCCCTGTCCGAAGGGGCAGTTTGGAGAAGGTTGTGCCA 1845

Qy 1619 AGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCC 1678
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1846 GTCGCTGTGACTGTGACCACTCTGATGGCTGTGACCCTGTTTATGGACGCTGTGAGTACC 1905

Qy 1679 TCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCA 1738
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1906 AGGCTGGCTGGATGGGTGCCCGCTGCCACCTGTCTGCCCTGAGGGCTTATGGGGAGTCA 1965

Qy 1739 ACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCT 1798
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1966 ACTGTAGCAACACCTGCACCTGCAAGAATGGGGGCACCTGTCTCCCTGAGAATGGCAACT 2025

Qy 1799 GCGAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTT 1858
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2026 GCGTGTGTGCACCCGATTCCGGGGCCCCCTCCTGCCAGAGATCCTGTGACGCTGGCCGCT 2085

Qy 1859 ATGGGCATCGCTGCAGCCAGACATGCCACAGTGCCTTACAGCAGCGGGGCCCTGCCACC 1918
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2086 ATGGCAAACGCTGTGTGC-----CCTGCAAGTGCCTAACCCTCC---TTCTGCCACC 2136

Qy 1919 ACATCACCGGCCTGTGTGACTGCTTGCCTGGCTTACAGGCGCCCTCTGCAATGAAGTGT 1978
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2137 CCTCGAACGGGACCTGCTACTGCCTGGCTGGCTGGACAGGCCCCGACTGCTCCCAGCGCT 2196

Qy 1979 GTCCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAA 2038

Db 2197 GCCCTCTGGGGACATTTGGTGCTAACTGCTCCCAGCCATGCCAGTGTGGTCTGGAGAAA 2256
 Qy 2039 CCTGTAACCCCATTTGACAGATCTTGTCACTGTTACCCCGGTTGGATTGG 2087
 Db 2257 AGTGCCACCCAGAGACTGGGGCTGTGTATGTCCCCAGGGCACAGTGG 2305

RATORED

DEFINITION *Rattus norvegicus* (clone REM4) ORF mRNA, partial cds.

ACCESSION L41686

VERSION L41686.1 GI:780366

KEYWORDS monoclonal autoantibody.

SOURCE *Rattus norvegicus* (Norway rat)

ORGANISM *Rattus norvegicus*

Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 660)

AUTHORS Asakura, K., Poqulis, R.J., Pease, L.R. and Rodriguez, M.

JOURNAL J. Neuroimmunol. 65 (1), 11-19 (1996)

MEDLINE 96235155

PUBMED 8642059

COMMENT Original source text: Rattus norvegicus (strain Holzman) cDNA to mRNA.

FEATURES	Location/Qualifiers
----------	---------------------

source

/organi

```
/mol type="mRNA"
```

```
/strain="Holzman
```

```
/db xref="taxon:1
```

```
/clone="REM#4"
```

```
/tissue type="brain"
```

```
/dev stage="neonate"
```

mRNA <1. .>660

CDS <1. .>660

```
/note="ORF"
```

```
/codon start=1
```

```
/protein_id="AAB05844.1"
```

```
/db xref="GI:780367"
```

/translation="RGATCDHITGECRCSPGYTGAFCEDLCPPGKHGPQCEQRCPCQN

GGVCHHVTGECSCPSGWMGTVCGOPCPEGRFGKNCSECOCHNGGACDAATGQCHCSP

GYTGFBCODECPVGTYGVRCAETCRCVNGGKCYHVS GTCLCEAGFSGEFCEARLCPEG

LYGIKCDKBCPCHLDNTHSCHPMMSGECGCKPGWSGLYCNETCSPGFYGEACQOICSCQ

NG"

ORIGIN

Best Local Similarity 89.6%; Pred. No. 8.7e-158;

Matches 589; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Ov 598 GGAGCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCC 657

```

      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      4 GGAGCGACCTGTGACCACATCACTGGGGAATGTCGTTGTTACCTGGATACACTGGAGCC 63

Qy      658 TTCTGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCT 717
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      64 TTCTGTGAGGATCTGTGTCCTCCTGGCAAACATGGCCACAGTGTGAGCAGAGATGTCCC 123

Qy      718 TGTCAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGG 777
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      124 TGCCAAAATGGAGGCGTGTGCCATCATGTCACTGGAGAGTGTCTCTGCCCTTCTGGCTGG 183

Qy      778 ATGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAA 837
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      184 ATGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAA 243

Qy      838 GAATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGT 897
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      244 GAATGTCAGTGCCACAATGGAGGAGCGTGTGATGCTGCCACAGGCCAATGTCACTGCAGC 303

Qy      898 CCAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTC 957
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      304 CCAGGATACACAGGGGAACGGTGTCAGGATGAATGTCTGTTGGAACCTATGGAGTTCGC 363

Qy      958 TGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCATGC 1017
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      364 TGTGCTGAGACCTGCAGGTGTGTTAATGGAGGGAAGTGTACCATGTGAGCGGCACATGC 423

Qy      1018 CTCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGCTC 1077
      || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      424 CTATGCGAAGCAGGTTTTTCAGGTGAATTTTGCAGGCTCGTCTGTGCCCTGAGGGGCTT 483

Qy      1078 TACGGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTCAC 1137
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      484 TACGGCATCAAATGTGACAAGCGGTGCCCTGCCACCTGGACAATACTCATAGCTGTCAT 543

Qy      1138 CCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACA 1197
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      544 CCCATGTCTGGAGAGTGTGGCTGCAAGCCGGGTTGGTCTGGACTGTACTGTAATGAAACA 603

Qy      1198 TGTTCCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGG 1254
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      604 TGCTCTCCTGGATTCTACGGGGAAGCTTGCCAACAGATCTGCAGCTGCCAAAACGGG 660

```

RESULT 15

AX079681

LOCUS AX079681 632 bp DNA linear PAT 22-FEB-2001

DEFINITION Sequence 425 from Patent WO0107611.

ACCESSION AX079681

VERSION AX079681.1 GI:13159250

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Baker,K.P., Goddard,A. and Wood,W.I.

TITLE Human polypeptides and methods for the use thereof
JOURNAL Patent: WO 0107611-A 425 01-FEB-2001;
Genentech, Inc. (US)

FEATURES Location/Qualifiers
source 1. .632
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 15.6%; Score 534.8; DB 6; Length 632;
Best Local Similarity 95.3%; Pred. No. 1.3e-153;
Matches 593; Conservative 0; Mismatches 24; Indels 5; Gaps 4;

```
Qy      7 ATTTCTTTGAACTCATGCCTGAGCTTTATTT--GTTTATTGTTATG-CCACTGGATTGGG 63
      |||| || || | ||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 ATTTTTTTGAAATTAATGCNTGAGCTTTATTTTGTTTAATTGTTATGCCCACTGGATTGGG 60

Qy      64 AC-AGCATCACCTCTGAATCTTGAAGACCCCTAATGTGTGTAGCCACTGGGAAAGCTACTC 122
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 ACAAGCATCACCTCTGAATTTTGAAGACCTTAATGTGTGTAGCCATTGNAAAGCTACTC 120

Qy      123 AGTGA CTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGCAC 182
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      121 AAGTGCTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGCAC 180

Qy      183 TGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTATCG 242
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      181 TGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTATCG 240

Qy      243 ACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAAAG 302
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      241 ACATGGGGAGAAGACTATG-ACAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAAAG 299

Qy      303 CGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCTCC 362
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      300 CGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCTCC 359

Qy      363 AAACACCTGTCAAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGATGG 422
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      360 AAACACCTGTCAAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGATGG 419

Qy      423 TGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGCAA 482
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      420 TGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGCAA 479

Qy      483 CCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCG 542
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      480 CCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCG 539
Qy      543 CTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGAGC 602
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      540 CTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAATGCCAGAATGGAGC 599
Qy      603 CACCTGCGACCACGTACGGGG 624
      ||||| ||||| ||||| |||||
Db      600 CACCTGCGACCACATCACGGGG 621
```

Search completed: March 30, 2004, 06:14:15
Job time : 13036 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2004, 23:13:51 ; Search time 1266 Seconds
(without alignments)
11486.242 Million cell updates/sec

Title: US-10-092-390-1
Perfect score: 3423
Sequence: 1 atgggttattttctttgaactc.....gcagcagcagcagtgatga 3423

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	3423	100.0	3423	6	AAD46318	Aad46318 Human EGF
2	3421.4	100.0	7522	9	ADE71243	Ade71243 Novel hum
3	3419.8	99.9	7522	9	ADD18689	Add18689 Human dis
4	1760	51.4	1761	6	AAD46319	Aad46319 Human EGF
5	1625.4	47.5	2909	5	AAS72220	Aas72220 DNA encod
6	1425.8	41.7	1448	2	AAX19959	Aax19959 Human Tan
7	1049.4	30.7	1273	7	ABZ36212	Abz36212 Human sec

	8	618.4	18.1	1074	5	AAS91826	Aas91826	DNA encod
	9	608.4	17.8	3567	4	AAF27791	Aaf27791	Rat TANGO
	10	600.6	17.5	3574	6	ABA00059	Aba00059	CADHP-6 c
	11	581.6	17.0	1402	5	AAS86746	Aas86746	DNA encod
	12	563.8	16.5	936	5	AAS72218	Aas72218	DNA encod
	13	563.8	16.5	936	7	ACD05889	Acd05889	Novel hum
	14	563.8	16.5	936	7	ACD05589	Acd05589	cDNA enco
	15	554	16.2	2295	5	AAS72219	Aas72219	DNA encod
	16	554	16.2	2295	5	AAS91825	Aas91825	DNA encod
	17	540.6	15.8	1970	4	AAH34884	Aah34884	Human col
	18	534.8	15.6	632	5	AAF93604	Aaf93604	cDNA isol
	19	528.6	15.4	3150	4	AAF27788	Aaf27788	Human TAN
	20	528.6	15.4	5036	4	AAF27787	Aaf27787	Human TAN
	21	476.4	13.9	1908	4	AAF27792	Aaf27792	Rat TANGO
	22	427.4	12.5	2220	7	ABX08773	Abx08773	Angiogene
	23	424.8	12.4	1205	4	AAD02807	Aad02807	HTLIH44 c
	24	420.2	12.3	3887	7	ABX34750	Abx34750	Human mdd
	25	396.8	11.6	7334	7	ABT33368	Abt33368	NOVX DNA
	26	395.4	11.6	2700	4	AAD02810	Aad02810	HFICU08 c
	27	393	11.5	4835	7	ABT33369	Abt33369	NOVX DNA
	28	390	11.4	4508	9	ADD78266	Add78266	Human CGD
	29	386.8	11.3	4733	7	ABT33366	Abt33366	NOVX DNA
c	30	386	11.3	3005	6	ABL53693	Abl53693	Human ubi
	31	383.8	11.2	7319	7	ABT33365	Abt33365	NOVX DNA
	32	383.8	11.2	7337	7	ABT33364	Abt33364	NOVX DNA
	33	377.6	11.0	390	5	AAS91824	Aas91824	DNA encod
	34	377.6	11.0	390	5	AAS86742	Aas86742	DNA encod
	35	371	10.8	412	5	AAS91823	Aas91823	DNA encod
	36	371	10.8	412	5	AAS86741	Aas86741	DNA encod
	37	358.4	10.5	5049	8	ADA21192	Ada21192	Human sec
	38	326.6	9.5	5096	8	ADA21191	Ada21191	Human sec
	39	315.4	9.2	1316	4	ABL19879	Abl19879	Drosophil
	40	272	7.9	353	3	AAA41581	Aaa41581	Human sec
c	41	243	7.1	3493	4	ABL19878	Abl19878	Drosophil
c	42	243	7.1	17478	4	ABL04034	Abl04034	Drosophil
	43	202.4	5.9	230	3	AAA44616	Aaa44616	Human sec
	44	178.2	5.2	971	5	AAS72263	Aas72263	DNA encod
	45	175	5.1	1686	4	AAD02811	Aad02811	HMEIR04 c

ALIGNMENTS

RESULT 1

AAD46318

ID AAD46318 standard; cDNA; 3423 BP.

XX

AC AAD46318;

XX

DT 27-JAN-2003 (first entry)

XX

DE Human EGF-family protein encoding cDNA #1.

XX

KW Human; EGF-family protein; novel human protein; NHP; drug discovery;

KW restriction fragment length polymorphism analysis; forensic biology;

KW toxicity; infectious disease; biological disorder; medical disorder;

KW mental disorder; gene therapy; gene; ss.

```

XX      Homo sapiens.
XX
FH      Key                Location/Qualifiers
FT      CDS                1. .3423
FT                        /*tag=  a
FT                        /product= "Human EGF-family protein #1"
XX
PN      WO200272611-A2.
XX
PD      19-SEP-2002.
XX
PF      06-MAR-2002; 2002WO-US007477.
XX
PR      12-MAR-2001; 2001US-0275013P.
XX
PA      (LEXI-) LEXICON GENETICS INC.
XX
PI      Yu X,  Miranda M;
XX
DR      WPI; 2002-723315/78.
DR      P-PSDB; AAE27985.
XX
PT      New novel human nucleic acids useful for e.g. identifying protein coding
PT      sequences and mapping unique genes to a particular chromosome, as DNA
PT      markers for restriction fragment length polymorphism analysis, or in
PT      forensic biology.
XX
PS      Claim 1; Page 36-37; 42pp; English.
XX
CC      The present sequence is a cDNA encoding human EGF-family protein, a novel
CC      human protein (NHP). The NHP sequences are useful for mapping unique
CC      genes to a particular chromosome; as DNA markers for restriction fragment
CC      length polymorphism analysis; in forensic biology; in defining and
CC      monitoring both drug action and toxicity; in identifying, selecting and
CC      validating novel molecular targets for drug discovery; in microarrays or
CC      other assay formats to screen collections of genetic material from
CC      patients who have a particular medical condition. The NHP peptides,
CC      fusion proteins, antibodies, antagonists and agonists can be used for
CC      detecting mutant NHPs or inappropriately expressed NHPs for the diagnosis
CC      of disease; for screening drugs for treatment of symptomatic or
CC      phenotypic manifestations of perturbing the normal function of NHP in the
CC      body and to treat diseases including infectious, mental, biological, or
CC      medical diseases or disorders. They are also used in gene therapy
XX
SQ      Sequence 3423 BP; 810 A; 904 C; 925 G; 784 T; 0 U; 0 Other;

```

Db	61	GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC	120
Qy	121	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	180
Db	121	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	180
Qy	181	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	240
Db	181	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	240
Qy	241	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	300
Db	241	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	300
Qy	301	AGCGGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	360
Db	301	AGCGGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	360
Qy	361	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	420
Db	361	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	420
Qy	421	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	480
Db	421	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	480
Qy	481	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	540
Db	481	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	540
Qy	541	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	600
Db	541	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	600
Qy	601	GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	660
Db	601	GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	660
Qy	661	TGTGAGGATCTTTGTCTCCTCGGTAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Db	661	TGTGAGGATCTTTGTCTCCTCGGTAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Qy	721	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Db	721	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Qy	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Db	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Qy	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Db	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Qy	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960
Db	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960

Qy	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1020
Db	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1020
Qy	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTAC	1080
Db	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTAC	1080
Qy	1081	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTCACCCC	1140
Db	1081	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTCACCCC	1140
Qy	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Db	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Qy	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Db	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Qy	1261	TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Db	1261	TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Qy	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1380
Db	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1380
Qy	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTG	1440
Db	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTG	1440
Qy	1441	GACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1500
Db	1441	GACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1500
Qy	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1560
Db	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1560
Qy	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1620
Db	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1620
Qy	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1680
Db	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1680
Qy	1681	CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1740
Db	1681	CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1740
Qy	1741	TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGC	1800
Db	1741	TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGC	1800

Qy	1801	GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTCTCAGAGGATCTGCTCCCCTGGTTTTTAT	1860
Db	1801	GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTCTCAGAGGATCTGCTCCCCTGGTTTTTAT	1860
Qy	1861	GGGCATCGCTGCAGCCAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCAC	1920
Db	1861	GGGCATCGCTGCAGCCAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCAC	1920
Qy	1921	ATCACCGGCCTGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT	1980
Db	1921	ATCACCGGCCTGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT	1980
Qy	1981	CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC	2040
Db	1981	CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC	2040
Qy	2041	TGTAACCCCATTTGACAGATCTTGTCTCAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTCT	2100
Db	2041	TGTAACCCCATTTGACAGATCTTGTCTCAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTCT	2100
Qy	2101	CAACCATGTCCACCTGCCCACTGGGGCCCAACTGCATCCACACGTGCAACTGCCATAAT	2160
Db	2101	CAACCATGTCCACCTGCCCACTGGGGCCCAACTGCATCCACACGTGCAACTGCCATAAT	2160
Qy	2161	GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC	2220
Db	2161	GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC	2220
Qy	2221	TACTGCACTCAGAGATGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAA	2280
Db	2221	TACTGCACTCAGAGATGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAA	2280
Qy	2281	TGTCAAAACGGAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTC	2340
Db	2281	TGTCAAAACGGAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTC	2340
Qy	2341	ATGGGACGGCACTGTGAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTCGCCAG	2400
Db	2341	ATGGGACGGCACTGTGAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTCGCCAG	2400
Qy	2401	ATATGTGATTGTCTGAACAACCTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC	2460
Db	2401	ATATGTGATTGTCTGAACAACCTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC	2460
Qy	2461	CCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGTTCATAGTTGGAAATCTGAAC	2520
Db	2461	CCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGTTCATAGTTGGAAATCTGAAC	2520
Qy	2521	AGCTTAAGCCGAACCACTGCTCTCCCTGCTGATTCCTACCAGATCGGGGCCATTGCA	2580
Db	2521	AGCTTAAGCCGAACCACTGCTCTCCCTGCTGATTCCTACCAGATCGGGGCCATTGCA	2580
Qy	2581	GGCATCATCATTTCTGTCTTAGTTGTTCTCTTCCCTACTGGCATTGTTTATTATATAGA	2640
Db	2581	GGCATCATCATTTCTGTCTTAGTTGTTCTCTTCCCTACTGGCATTGTTTATTATATAGA	2640
Qy	2641	CACAAGCAGAAGGGAAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2700

Db	2641	 CACAAGCAGAAAGGAAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2700
Qy	2701	GTCGTCAATGCAGATTATACCATTTTCAGGAACCCTTCCTCACAGCAATGGTGGAAACGCT	2760
Db	2701	 GTCGTCAATGCAGATTATACCATTTTCAGGAACCCTTCCTCACAGCAATGGTGGAAACGCT	2760
Qy	2761	AATAGCCACTACTTCACCAATCCCAGTTACCACACGCTCACCCAGTGTGCCACATCCCCCT	2820
Db	2761	 AATAGCCACTACTTCACCAATCCCAGTTACCACACGCTCACCCAGTGTGCCACATCCCCCT	2820
Qy	2821	CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	2880
Db	2821	 CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	2880
Qy	2881	AATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	2940
Db	2881	 AATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	2940
Qy	2941	CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3000
Db	2941	 CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3000
Qy	3001	AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAC	3060
Db	3001	 AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAC	3060
Qy	3061	TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3120
Db	3061	 TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3120
Qy	3121	CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCA	3180
Db	3121	 CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCA	3180
Qy	3181	TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA	3240
Db	3181	 TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA	3240
Qy	3241	GTGAGTGTGTGTCGAAGGAGTATTCAGCAATAATGGGCGTCTCTCCAGGATCCATATGAC	3300
Db	3241	 GTGAGTGTGTGTCGAAGGAGTATTCAGCAATAATGGGCGTCTCTCCAGGATCCATATGAC	3300
Qy	3301	CTCCCAAAGAACAGTCACATCCCTTGTCATTATGACCTGCTGCCAGTCCGAGACAGTTCA	3360
Db	3301	 CTCCCAAAGAACAGTCACATCCCTTGTCATTATGACCTGCTGCCAGTCCGAGACAGTTCA	3360
Qy	3361	TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGTGA	3420
Db	3361	 TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGTGA	3420
Qy	3421	TGA 3423	
Db	3421	 TGA 3423	

ADE71243

ID ADE71243 standard; DNA; 7522 BP.

XX

AC ADE71243;

XX

DT 29-JAN-2004 (first entry)

XX

DE Novel human protein coding sequence #59.

XX

KW human; novel protein; drug; gene; ds.

XX

OS Homo sapiens.

XX

PN JP2002345493-A.

XX

PD 03-DEC-2002.

XX

PF 29-MAR-2001; 2002JP-00049046.

XX

PR 29-MAR-2001; 2001JP-00095524.

XX

PA (KAZU-) ZH KAZUSA DNA KENKYUSHO.

XX

DR WPI; 2003-460885/44.

DR P-PSDB; ADE71305.

XX

PT A gene and a protein encoded by it, used in drugs.

XX

PS Claim 1; SEQ ID NO 60; 257pp; Japanese.

XX

CC The invention comprises the amino acid and coding sequences of novel
CC human proteins. The DNA and protein sequences of the invention are used
CC in drugs. The present DNA sequence encodes a novel human protein of the
CC invention.

XX

SQ Sequence 7522 BP; 2130 A; 1554 C; 1697 G; 2141 T; 0 U; 0 Other;

Query Match 100.0%; Score 3421.4; DB 9; Length 7522;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT	60
Db	204	ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT	263
Qy	61	GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC	120
Db	264	GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC	323
Qy	121	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	180
Db	324	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	383
Qy	181	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	240
Db	384	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	443

Qy	241	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	300
Db	444	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	503
Qy	301	AGCGGGGAAATGTGTGTCCTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	360
Db	504	AGCGGGGAAATGTGTGTCCTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	563
Qy	361	CCAAACACCTGTCAGTGTCAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	420
Db	564	CCAAACACCTGTCAGTGTCAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	623
Qy	421	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	480
Db	624	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	683
Qy	481	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	540
Db	684	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	743
Qy	541	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	600
Db	744	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	803
Qy	601	GCCACCTGCGACCACGTCACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	660
Db	804	GCCACCTGCGACCACGTCACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	863
Qy	661	TGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Db	864	TGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	923
Qy	721	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Db	924	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	983
Qy	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Db	984	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	1043
Qy	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Db	1044	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	1103
Qy	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960
Db	1104	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	1163
Qy	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTTACCACGTGAGCGGCGCATGCCTC	1020
Db	1164	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTTACCACGTGAGCGGCGCATGCCTC	1223
Qy	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCCAAGCACGCCTGTGTCCTGAGGGGCTCTAC	1080
Db	1224	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCCAAGCACGCCTGTGTCCTGAGGGGCTCTAC	1283
Qy	1081	GGCATCAAATGTGACAAACGGTGTCCTGCCACTTGGAACAACTCATAGCTGTCACCCC	1140

Db	1284	 GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTACCCC	1343
Qy	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Db	1344	 ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1403
Qy	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Db	1404	 TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1463
Qy	1261	TGTGACAGTGTGACTGGAAAGTGACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Db	1464	 TGTGACAGTGTGACTGGAAAGTGACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1523
Qy	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1380
Db	1524	 ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1583
Qy	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTAATTGCAAGGCAGGCTGGCACGGGGTG	1440
Db	1584	 GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTAATTGCAAGGCAGGCTGGCACGGGGTG	1643
Qy	1441	GACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1500
Db	1644	 GACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1703
Qy	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1560
Db	1704	 TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1763
Qy	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1620
Db	1764	 CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1823
Qy	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1680
Db	1824	 CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1883
Qy	1681	CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1740
Db	1884	 CCCGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1943
Qy	1741	TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGC	1800
Db	1944	 TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGC	2003
Qy	1801	GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTCAGAGGATCTGCTCCCCTGGTTTTTAT	1860
Db	2004	 GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTCAGAGGATCTGCTCCCCTGGTTTTTAT	2063
Qy	1861	GGGCATCGCTGCAGCCAGACATGCCACAGTGC GTTCACAGCAGCGGGCCCTGCCACCAC	1920
Db	2064	 GGGCATCGCTGCAGCCAGACATGCCACAGTGC GTTCACAGCAGCGGGCCCTGCCACCAC	2123
Qy	1921	ATCACCGGCCTGTGTGACTGCTTGCCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT	1980

Db	2124	ATCACCGGCCTGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT	2183
Qy	1981	CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC	2040
Db	2184	CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC	2243
Qy	2041	TGTAACCCCATTTGACAGATCTTGTCAAGTGTACCCCGTTGGATTGGCAGTGAAGTGTCT	2100
Db	2244	TGTAACCCCATTTGACAGATCTTGTCAAGTGTACCCCGTTGGATTGGCAGTGAAGTGTCT	2303
Qy	2101	CAACCATGTCCACCTGCCCCTGGGGCCCAACTGCATCCACACGTGCAACTGCCATAAT	2160
Db	2304	CAACCATGTCCACCTGCCCCTGGGGCCCAACTGCATCCACACGTGCAACTGCCATAAT	2363
Qy	2161	GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC	2220
Db	2364	GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC	2423
Qy	2221	TACTGCACTCAGAGATGTCCTCTAGGGTTTATGGAAAAGATTGTGCACTGATATGCCAA	2280
Db	2424	TACTGCACTCAGAGATGTCCTCTAGGGTTTATGGAAAAGATTGTGCACTGATATGCCAA	2483
Qy	2281	TGTCAAAACGGAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTC	2340
Db	2484	TGTCAAAACGGAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTC	2543
Qy	2341	ATGGGACGGCACTGTGAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAG	2400
Db	2544	ATGGGACGGCACTGTGAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAG	2603
Qy	2401	ATATGTGATTGTCTGAACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC	2460
Db	2604	ATATGTGATTGTCTGAACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC	2663
Qy	2461	CCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAAC	2520
Db	2664	CCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAAC	2723
Qy	2521	AGCTTAAGCCGAACCACTGCTCTCCCTGCTGATTCCTACCAGATCGGGGCCATTGCA	2580
Db	2724	AGCTTAAGCCGAACCACTGCTCTCCCTGCTGATTCCTACCAGATCGGGGCCATTGCA	2783
Qy	2581	GGCATCATCATTCTTGTCTAGTTGTTCTCTTCCTACTGGCATTGTTTATTATTTATAGA	2640
Db	2784	GGCATCATCATTCTTGTCTAGTTGTTCTCTTCCTACTGGCATTGTTTATTATTTATAGA	2843
Qy	2641	CACAAGCAGAAGGGAAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2700
Db	2844	CACAAGCAGAAGGGAAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2903
Qy	2701	GTCGTCAATGCAGATTATACCATTTTTCCTCACAGCAATGGTGGAAACGCT	2760
Db	2904	GTCGTCAATGCAGATTATACCATTTTTCCTCACAGCAATGGTGGAAACGCT	2963
Qy	2761	AATAGCCACTACTTCACCAATCCAGTTACCACACGCTCACCAGTGTGCCACATCCCCCT	2820
Db	2964	AATAGCCACTACTTCACCAATCCAGTTACCACACGCTCACCAGTGTGCCACATCCCCCT	3023

Qy	2821	CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	2880
Db	3024	CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	3083
Qy	2881	AATCTTAAAAATGTGAACCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	2940
Db	3084	AATCTTAAAAATGTGAACCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	3143
Qy	2941	CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3000
Db	3144	CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3203
Qy	3001	AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAC	3060
Db	3204	AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAC	3263
Qy	3061	TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3120
Db	3264	TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3323
Qy	3121	CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCA	3180
Db	3324	CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCA	3383
Qy	3181	TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA	3240
Db	3384	TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA	3443
Qy	3241	GTGAGTGTTGTCCAAGGAGTATTCAGCAATAATGGGCGTCTCTCCCAGGATCCATATGAC	3300
Db	3444	GTGAGTGTTGTCCAAGGAGTATTCAGCAATAATGGGCGTCTCTCCCAGGATCCATATGAC	3503
Qy	3301	CTCCCAAAGAACAGTCACATCCCTTGTCATTATGACCTGCTGCCAGTCCGAGACAGTTCA	3360
Db	3504	CTCCCAAAGAACAGTCACATCCCTTGTCATTATGACCTGCTGCCAGTCCGAGACAGTTCA	3563
Qy	3361	TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGTGAA	3420
Db	3564	TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGTGAA	3623
Qy	3421	TGA	3423
Db	3624	TGA	3626

RESULT 3

ADD18689

ID ADD18689 standard; DNA; 7522 BP.

XX

AC ADD18689;

XX

DT 15-JAN-2004 (first entry)

XX

DE Human disease related protein DNA sequence SeqID120.

XX

KW human; disease state; cytostatic; antiinflammatory; ophthalmological;

KW antiarteriosclerotic; vulnerary; gene therapy;
 KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
 KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
 KW glucose transportation; catecholamine synthesis; iron transport;
 KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
 KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
 KW inflammatory condition; wound healing; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003018621-A2.
 XX
 PD 06-MAR-2003.
 XX
 PF 23-AUG-2002; 2002WO-GB003892.
 XX
 PR 23-AUG-2001; 2001GB-00020558.
 PR 05-OCT-2001; 2001GB-00024037.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
 XX
 DR WPI; 2003-290046/28.
 DR P-PSDB; ADD18688.
 XX
 PT New substantially purified polypeptide, useful for diagnosing or treating
 PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
 PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
 PT wound healing.
 XX
 PS Claim 27; SEQ ID NO 120; 424pp; English.
 XX
 CC This invention relates to novel human genes and gene product which are
 CC implicated in certain disease states. Compounds which modulate the
 CC proteins of the invention may have cytostatic, antiinflammatory,
 CC ophthalmological, antiarteriosclerotic or vulnerary activities. The
 CC sequences of the invention may be useful for gene therapy. The invention
 CC may be useful for diagnosing or treating a hypoxia-regulated condition,
 CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,
 CC erythropoiesis, or the biological response to hypoxia conditions
 CC including processes such as glycolysis, gluconeogenesis, glucose
 CC transportation, catecholamine synthesis, iron transport or nitric oxide
 CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
 CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
 CC inflammatory conditions or wound healing. The present sequence is that of
 CC a disease related protein encoding DNA sequence of the invention.
 XX
 SQ Sequence 7522 BP; 2130 A; 1555 C; 1696 G; 2141 T; 0 U; 0 Other;

Query Match 99.9%; Score 3419.8; DB 9; Length 7522;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3421; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 204 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 263

Qy	61	GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC	120
Db	264	GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC	323
Qy	121	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	180
Db	324	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	383
Qy	181	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	240
Db	384	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	443
Qy	241	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	300
Db	444	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	503
Qy	301	AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	360
Db	504	AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	563
Qy	361	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	420
Db	564	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	623
Qy	421	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	480
Db	624	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	683
Qy	481	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	540
Db	684	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	743
Qy	541	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	600
Db	744	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	803
Qy	601	GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	660
Db	804	GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	863
Qy	661	TGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Db	864	TGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	923
Qy	721	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Db	924	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	983
Qy	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAAGTGTCCCAAGAA	840
Db	984	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAAGTGTCCCAAGAA	1043
Qy	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Db	1044	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	1103

Qy	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTGGGACCTATGGCGTTCTCTGT	960
Db	1104	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTGGGACCTATGGCGTTCTCTGT	1163
Qy	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1020
Db	1164	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCCCATGCCTC	1223
Qy	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTAC	1080
Db	1224	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTAC	1283
Qy	1081	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGAAAACTCATAGCTGTCACCCC	1140
Db	1284	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGAAAACTCATAGCTGTCACCCC	1343
Qy	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Db	1344	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1403
Qy	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Db	1404	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1463
Qy	1261	TGTGACAGTGTGACTGGAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Db	1464	TGTGACAGTGTGACTGGAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1523
Qy	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAAT	1380
Db	1524	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAAT	1583
Qy	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTAATTGCAAGGCAGGCTGGCACGGGGTG	1440
Db	1584	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTAATTGCAAGGCAGGCTGGCACGGGGTG	1643
Qy	1441	GACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1500
Db	1644	GACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1703
Qy	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1560
Db	1704	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1763
Qy	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1620
Db	1764	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1823
Qy	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGGCCATTGCCGCTGCCTC	1680
Db	1824	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGGCCATTGCCGCTGCCTC	1883
Qy	1681	CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1740
Db	1884	CCCGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1943
Qy	1741	TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCTGATGATGGCATCTGC	1800

Db	1944	 TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGC	2003
Qy	1801	GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGAGGATCTGCTCCCCTGGTTTTAT	1860
Db	2004	 GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGAGGATCTGCTCCCCTGGTTTTAT	2063
Qy	1861	GGGCATCGCTGCAGCCAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCAC	1920
Db	2064	 GGGCATCGCTGCAGCCAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCAC	2123
Qy	1921	ATCACCGGCCTGTGTGACTGCTTGCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT	1980
Db	2124	 ATCACCGGCCTGTGTGACTGCTTGCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT	2183
Qy	1981	CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC	2040
Db	2184	 CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC	2243
Qy	2041	TGTAACCCCATTGACAGATCTTGTCAGTGTTACCCCGGTTGGATTGGCAGTGACTGCTCT	2100
Db	2244	 TGTAACCCCATTGACAGATCTTGTCAGTGTTACCCCGGTTGGATTGGCAGTGACTGCTCT	2303
Qy	2101	CAACCATGTCCACCTGCCACTGGGGCCCAAAGTGCATCCACACGTGCAACTGCCATAAT	2160
Db	2304	 CAACCATGTCCACCTGCCACTGGGGCCCAAAGTGCATCCACACGTGCAACTGCCATAAT	2363
Qy	2161	GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC	2220
Db	2364	 GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC	2423
Qy	2221	TACTGCACTCAGAGATGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAA	2280
Db	2424	 TACTGCACTCAGAGATGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAA	2483
Qy	2281	TGTCAAACGGAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTC	2340
Db	2484	 TGTCAAACGGAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTC	2543
Qy	2341	ATGGGACGGCACTGTGAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTCGCCAG	2400
Db	2544	 ATGGGACGGCACTGTGAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTCGCCAG	2603
Qy	2401	ATATGTGATTGTCTGAACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC	2460
Db	2604	 ATATGTGATTGTCTGAACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC	2663
Qy	2461	CCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAAC	2520
Db	2664	 CCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAAC	2723
Qy	2521	AGCTTAAGCCGAACCACTGCTCTCCCTGCTGATTTCCTACCAGATCGGGGCCATTGCA	2580
Db	2724	 AGCTTAAGCCGAACCACTGCTCTCCCTGCTGATTTCCTACCAGATCGGGGCCATTGCA	2783
Qy	2581	GGCATCATCATTCTTGTCTAGTTGTTCTCTTCCTACTGGCATTGTTTATTATAGA	2640

Db	2784	GGCATCATCATTCTTGTCCTAGTTGTTCTCTTCCCTACTGGCATTGTTTCATTATTTATAGA	2843
Qy	2641	CACAAGCAGAAGGGAAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2700
Db	2844	CACAAGCAGAAGGGAAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2903
Qy	2701	GTCGTCAATGCAGATTATACCATTTTCAGGAACCCTTCCTCACAGCAATGGTGGAAACGCT	2760
Db	2904	GTCGTCAATGCAGATTATACCATTTTCAGGAACCCTTCCTCACAGCAATGGTGGAAACGCT	2963
Qy	2761	AATAGCCACTACTTCACCAATCCCAGTTACCACACGCTCACCCAGTGTGCCACATCCCCT	2820
Db	2964	AATAGCCACTACTTCACCAATCCCAGTTACCACACGCTCACCCAGTGTGCCACATCCCCT	3023
Qy	2821	CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	2880
Db	3024	CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	3083
Qy	2881	AATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	2940
Db	3084	AATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	3143
Qy	2941	CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3000
Db	3144	CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3203
Qy	3001	AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAC	3060
Db	3204	AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAC	3263
Qy	3061	TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3120
Db	3264	TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3323
Qy	3121	CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCA	3180
Db	3324	CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCA	3383
Qy	3181	TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA	3240
Db	3384	TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA	3443
Qy	3241	GTGAGTGTTGTCCAAGGAGTATTAGCAATAATGGGCGTCTCTCCAGGATCCATATGAC	3300
Db	3444	GTGAGTGTTGTCCAAGGAGTATTAGCAATAATGGGCGTCTCTCCAGGATCCATATGAC	3503
Qy	3301	CTCCCAAAGAACAGTCACATCCCTTGTGATTATGACCTGCTGCCAGTCCGAGACAGTTCA	3360
Db	3504	CTCCCAAAGAACAGTCACATCCCTTGTGATTATGACCTGCTGCCAGTCCGAGACAGTTCA	3563
Qy	3361	TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGTGA	3420
Db	3564	TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGTGA	3623
Qy	3421	TGA	3423
Db	3624	TGA	3626

RESULT 4

AAD46319

ID AAD46319 standard; cDNA; 1761 BP.

XX

AC AAD46319;

XX

DT 27-JAN-2003 (first entry)

XX

DE Human EGF-family protein encoding cDNA #2.

XX

KW Human; EGF-family protein; novel human protein; NHP; drug discovery;

KW restriction fragment length polymorphism analysis; forensic biology;

KW toxicity; infectious disease; biological disorder; medical disorder;

KW mental disorder; gene therapy; gene; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1. .1761

FT /*tag= a

FT /product= "Human EGF-family protein #2"

XX

PN WO200272611-A2.

XX

PD 19-SEP-2002.

XX

PF 06-MAR-2002; 2002WO-US007477.

XX

PR 12-MAR-2001; 2001US-0275013P.

XX

PA (LEXI-) LEXICON GENETICS INC.

XX

PI Yu X, Miranda M;

XX

DR WPI; 2002-723315/78.

DR P-PSDB; AAE27986.

XX

PT New novel human nucleic acids useful for e.g. identifying protein coding

PT sequences and mapping unique genes to a particular chromosome, as DNA

PT markers for restriction fragment length polymorphism analysis, or in

PT forensic biology.

XX

PS Disclosure; Page 40; 42pp; English.

XX

CC The present sequence is a cDNA encoding human EGF-family protein, a novel

CC human protein (NHP). The NHP sequences are useful for mapping unique

CC genes to a particular chromosome; as DNA markers for restriction fragment

CC length polymorphism analysis; in forensic biology; in defining and

CC monitoring both drug action and toxicity; in identifying, selecting and

CC validating novel molecular targets for drug discovery; in microarrays or

CC other assay formats to screen collections of genetic material from

CC patients who have a particular medical condition. The NHP peptides,

CC fusion proteins, antibodies, antagonists and agonists can be used for

CC detecting mutant NHPs or inappropriately expressed NHPs for the diagnosis

CC of disease; for screening drugs for treatment of symptomatic or

CC phenotypic manifestations of perturbing the normal function of NHP in the
CC body and to treat diseases including infectious, mental, biological, or
CC medical diseases or disorders. They are also used in gene therapy
XX
SQ Sequence 1761 BP; 370 A; 465 C; 525 G; 401 T; 0 U; 0 Other;

Query Match 51.4%; Score 1760; DB 6; Length 1761;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 60
        |||
Db      1 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 60

Qy     61 GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC 120
        |||
Db     61 GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC 120

Qy    121 TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC 180
        |||
Db    121 TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC 180

Qy    181 ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT 240
        |||
Db    181 ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT 240

Qy    241 CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA 300
        |||
Db    241 CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA 300

Qy    301 AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT 360
        |||
Db    301 AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT 360

Qy    361 CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT 420
        |||
Db    361 CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT 420

Qy    421 GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC 480
        |||
Db    421 GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC 480

Qy    481 AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC 540
        |||
Db    481 AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC 540

Qy    541 CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA 600
        |||
Db    541 CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA 600

Qy    601 GCCACCTGCGACCACGTCACGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTC 660
        |||
Db    601 GCCACCTGCGACCACGTCACGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTC 660

Qy    661 TGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT 720
        |||
Db    661 TGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT 720
```

Qy	721	CAAAATGGAGGAGTGTGTATCACGTCCTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Db	721	CAAAATGGAGGAGTGTGTATCACGTCCTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Qy	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Db	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Qy	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Db	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Qy	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960
Db	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960
Qy	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1020
Db	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1020
Qy	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGCTCTAC	1080
Db	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGCTCTAC	1080
Qy	1081	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTCAACCC	1140
Db	1081	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTCAACCC	1140
Qy	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Db	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Qy	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAATGGGGCAGAC	1260
Db	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAATGGGGCAGAC	1260
Qy	1261	TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Db	1261	TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Qy	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1380
Db	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1380
Qy	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGACTTGCAAGGCAGGCTGGCACGGGGTG	1440
Db	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGACTTGCAAGGCAGGCTGGCACGGGGTG	1440
Qy	1441	GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1500
Db	1441	GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1500
Qy	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1560
Db	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1560

Qy	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGGCTGAACTGTGCTGAG	1620
Db	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGGCTGAACTGTGCTGAG	1620
Qy	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1680
Db	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1680
Qy	1681	CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1740
Db	1681	CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1740
Qy	1741	TGCTCCCTGCCCTGCTACTG	1760
Db	1741	TGCTCCCTGCCCTGCTACTG	1760

RESULT 5

AAS72220

ID AAS72220 standard; cDNA; 2909 BP.

XX

AC AAS72220;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #8024.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

DR P-PSDB; ABG08033.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX

PS Claim 1; SEQ ID NO 8024; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 2909 BP; 688 A; 777 C; 760 G; 684 T; 0 U; 0 Other;

Query Match 47.5%; Score 1625.4; DB 5; Length 2909;
 Best Local Similarity 80.2%; Pred. No. 0;
 Matches 2258; Conservative 0; Mismatches 36; Indels 521; Gaps 9;

Qy	1130	GCTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTA	1189
Db	95	GCTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTA	154
Qy	1190	ATGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAA	1249
Db	155	ATGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAA	214
Qy	1250	ATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAA	1309
Db	215	ATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAA	274
Qy	1310	TTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTG	1369
Db	275	TTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTG	334
Qy	1370	GCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGACTTGCAAGGCAGGCT	1429
Db	335	GCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGACTTGCAAGGCAGGCT	394
Qy	1430	GGCACGGGGTGGACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACT	1489
Db	395	GGCACGGGGTGGACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACT	454
Qy	1490	TAACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTG	1549
Db	455	TAACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTG	514
Qy	1550	CACCTGGATGGCGCGGGGAGAAATGCGAACTTCCTGCCAGGATGGCACGTACGGGCTGA	1609
Db	515	CACCTGGATGGCGCGGGGAGAAATGCGAACTTCCTGCCAGGATGGCACGTACGGGCTGA	574

Qy 1610 ACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGGCCATT 1669
 |||
 Db 575 ACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGGCCATT 634

Qy 1670 GCCGCTGCCTCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCT 1729
 |||
 Db 635 GCCGCTGCCTCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCT 694

Qy 1730 GGGGCCCCAACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATG 1789
 |||
 Db 695 GGGGCCCCAACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATG 754

Qy 1790 ATGGCATCTGCGAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCC 1849
 |||
 Db 755 ATGGCATCTGCGAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCC 814

Qy 1850 CTGGTTTTTATGGGCATCGCTGCAGCCAGACATGCCCACAGTGCCTTCACAGCAGCGGGC 1909
 |||
 Db 815 CTGGTTTTTATGGGCATCGCTGCAGCCAGACATGCCCACAGTGCCTTCACAGCAGCGGGC 874

Qy 1910 CCTGCCACCACATCACCGGCCTGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCA 1969
 |||
 Db 875 CCTGCCACCACATCACCGGCCTGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCA 934

Qy 1970 ATGA-----AGTGTGTCCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTG 2018
 |||
 Db 935 ATGAAGCTTATTCACAGTGTGTCCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTG 994

Qy 2019 TACCTGCACCAACAACGGAACCTGTAACCCCATTTGACAGATCTTGTGAGTGTACCCCGG 2078
 |||
 Db 995 TACCTGCACCAACAACGGAACCTGTAACCCCATTTGACAGATCTTGTGAGTGTACCCCGG 1054

Qy 2079 TTGGATTGGCAGTGACTGCTCTCAACCATGTCCACCTGCCCACTGGGGCCCAAACCTGCAT 2138
 |||
 Db 1055 TTGGATTGGCAGTGACTGCTCTCAACCATGTCCACCTGCCCACTGGGGCCCAAACCTGCAT 1114

Qy 2139 CCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATG 2198
 |||
 Db 1115 CCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATG 1174

Qy 2199 CACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGATGTCC----- 2240
 |||
 Db 1175 CACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAAGATCCCCAAGACACTCTTGCAGGG 1234

Qy 2241 -----TCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACG 2290
 |||
 Db 1235 CAGCTGCCAGCCCATGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACG 1294

Qy 2291 GAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGC 2350
 |||
 Db 1295 GAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGC 1354

Qy 2351 ACTGTGAGCAGA----- 2362
 |||
 Db 1355 ACTGTGAGCAGAAGGTCAGGCCTCCCTGGGATCATCGCTGGTTACTCACTGCTCTGGGCG 1414

Qy 2363 ----- 2362

Db 1415 GAGGAGGTGTGACTACAAGAATGAAGACAGAATTCAAGTTTCCATTCTGTTCTTTTGGG 1474

Qy 2363 ----- 2362

Db 1475 CTCTCCCTTCCTCTCCTTCTTATTTTTGGAATGTGGCTGCTCAGAGCCTTAAAAGATCCA 1534

Qy 2363 ----- 2362

Db 1535 GTCGAGCTTTCTTCATGGCAGAGGCAGAACCCGGGTCTCACATTGGAGGACAGTACATTC 1594

Qy 2363 ----- 2362

Db 1595 GGTGGGGAGGAGGGCTGGTGGCCCAGGGCCAATCCCTGTTGCTGCCCTGTGCAGTCTGGA 1654

Qy 2363 ----- 2362

Db 1655 CTGTGAGTGCCACCATGATTCCAGGAATGCTGTCCAGTTCTGGGACACTCTTGGGGGTAC 1714

Qy 2363 ----- 2362

Db 1715 AAGTCAGCCTCAATAGGAATCCCCTGAAAGGACTCAGCTCTAGATGTGCAGGGCTTGCAG 1774

Qy 2363 ----- 2362

Db 1775 TCAGAGACAGCCTTGCTCCAAATCCCAAGGTTGGAAAGCTACATTTGACTTTCTTCTC 1834

Qy 2363 ---AGTGCCCTTCAGGAACATATGGCTATGGCTGTGCCAGATATGTGATTGTCTGAACA 2419
 |||

Db 1835 TAGAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCCAGATATGTGATTGTCTGAACA 1894

Qy 2420 ACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGAGCGA 2479
 |||

Db 1895 ACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGAGCGA 1954

Qy 2480 GATGTGATCAAGCTGGTGTATCATAGTTGGAATCTGAACAGCTTAAGCCGAACCAGTA 2539
 |||

Db 1955 GATGTGATCAAGCTGGTGTATCATAGTTGGAATCTGAACAGCTTAAGCCGAACCAGTA 2014

Qy 2540 CTGCTCTCCCTGCTGATTCCCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTTGTCC 2599
 |||

Db 2015 CTGCTCTCCCTGCTGATTCCCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTTGTCC 2074

Qy 2600 TAGTTGTTCTCTTCCTACTGGCATTGTTCAATTATTTATAGACACAAGCAGAAGGGAAAGG 2659
 |||

Db 2075 TAGTTGTTCTCTTCCTACTGGCATTGTTCAATTATTTATAGACACAAGCAGAAGGGAAAGG 2134

Qy 2660 AATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGGGTCGTCAATGCAGATTATA 2719
 |||

Db 2135 AATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGGGTCGTCAATGCAGATTATA 2194

Qy 2720 CCATTTTCAGGAACCCCTTCCTCACAGCAATGGTGGAAACGCTAATAGCCACTACTTCACCA 2779
 |||

Db 2195 CCATTTTCAGGAACCCCTTCCTCACAGCAATGGTGGAAACGCTAATAGCCACTACTTCACCA 2254

Qy 2780 ATCCCAGTTACCACACGCTCACCCAGTGTGCCACATCCCCT---CACGTCAACAACAGGG 2836

Db	2255	ATCCCAGTTACACACGCTCACCCAGTGTGCCACATCCCCTTCACGTTCAACAACAGGGG	2314
Qy	2837	ACAGGATGACTG-TCACGAAGTCAAAAAACAATCAACTGTTTGTGAATCTTAAAAATGTG	2895
Db	2315	ACAGGATGACTGTTTCACGAGTTCAAAAACAATCAACTGTTTGTGAATCTTAAAAATGTG	2374
Qy	2896	AACCCTGGGAAGAGAGGCCCTGT-GGGGGACTGCACTGGGACATTGCCGGCTGACTGGAA	2954
Db	2375	AACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTGCCGGCTGACTGGAA	2434
Qy	2955	ACATGGCGGCTACCTCAACGAGCT--CGGTGCTTTTGGACTTGACAGAAGCTATAT-GGG	3011
Db	2435	ACATGGCGGCTACCTCAACGAGCTTCGGTGCTTTTGGGACTTGACAGAAGCTATTTGGGG	2494
Qy	3012	AAAATCCTTAAAAGACC---TGGGAAAGAATTCTGAATATAATTCAAGTAAGTCTCCCT	3068
Db	2495	AAAATCCTTAAAAGGACCTGGGGAAAGGATTTTGAATATAATTCAAGTAAGTCTCCCT	2554
Qy	3069	AAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATCCCGAAAAG	3128
Db	2555	AAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATCCCGAAAAG	2614
Qy	3129	CTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGGCACGAAGAGATTCCCATATGCAGA	3188
Db	2615	CTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGGCACGAAGAGATTCCCATATGCAGA	2674
Qy	3189	GATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACAGTGAGTGT	3248
Db	2675	GATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACAGTGAGTGT	2734
Qy	3249	TGTCCAAGGAGTATTCAGCAATAATGGGCGTCTCTCCAGGATCCATATGACCTCCCAA	3308
Db	2735	TGTCCAAGGAGTATTCAGCAATAATGGGCGTCTCTCCAGGATCCATATGACCTCCCAA	2794
Qy	3309	GAACAGTCACATCCCTTGTCATTATGACCTGCTGCCAGTCCGAGACAGTTCATCCTCCCC	3368
Db	2795	GAACAGTCACATCCCTTGTCATTATGACCTGCTGCCAGTCCGAGACAGTTCATCCTCCCC	2854
Qy	3369	TAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGCAGTGAATGA	3423
Db	2855	TAAGCAAGAGGACAGTGGTGGTAGCAGCAGCAACAGCAGCAGCAGCAGTGAATGA	2909

RESULT 6

AAX19959

ID AAX19959 standard; DNA; 1448 BP.

XX

AC AAX19959;

XX

DT 15-JUN-1999 (first entry)

XX

DE Human Tango-83 5'-portion.

XX

KW Human; Tango-71; Tango-73; Tango-74; Tango-76; Tango-83; diagnosis;

KW detection; ds.

XX

OS Homo sapiens.
XX
PN WO9907850-A1.
XX
PD 18-FEB-1999.
XX
PF 06-AUG-1998; 98WO-US016502.
XX
PR 06-AUG-1997; 97US-0054966P.
PR 05-SEP-1997; 97US-0058108P.
XX
PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI Holtzman DA, Goodearl ADJ;
XX
DR WPI; 1999-167426/14.
XX
PT New TANGO polypeptides and nucleic acids encoding them - useful as
PT diagnostic agents and for treating disorders caused by aberrant
PT expression of TANGO.
XX
PS Claim 1; Fig 7; 84pp; English.
XX
CC The present sequence is a 5'-portion of Tango-83. Tango polypeptides are
CC useful for identifying compounds which bind the polypeptide via direct
CC binding, competition binding assays or Tango-71, -73, -74, 76 or -83-
CC mediated signal transduction. Tango polypeptides are also useful for
CC identifying modulating compounds by determining effect on Tango activity.
CC Tango polypeptides and nucleic acids are useful for diagnosing diseases
CC related to aberrant expression of Tango, and Tango polypeptides are
CC useful for raising antibodies which can be used in diagnostic assays for
CC detection of Tango, and also for generating anti-idiotypic antibodies for
CC prevention and protection
XX
SQ Sequence 1448 BP; 308 A; 407 C; 400 G; 333 T; 0 U; 0 Other;

Query Match 41.7%; Score 1425.8; DB 2; Length 1448;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1427; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1216 GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT 1275
|||||
Db 18 GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT 77

Qy 1276 GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG 1335
|||||
Db 78 GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG 137

Qy 1336 GGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT 1395
|||||
Db 138 GGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT 197

Qy 1396 CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGA 1455
|||||
Db 198 CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGA 257

Qy 1456 TGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGA 1515

Db	258	TGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGA	317
Qy	1516	GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC	1575
Db	318	GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC	377
Qy	1576	GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACCTGTGCTGAGCGCTGCGACTGCAGC	1635
Db	378	GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACCTGTGCTGAGCGCTGCGACTGCAGC	437
Qy	1636	CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT	1695
Db	438	CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT	497
Qy	1696	GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC	1755
Db	498	GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC	557
Qy	1756	TACTGTAAAAATGGGGCTTCATGCTCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC	1815
Db	558	TACTGTAAAAATGGGGCTTCATGCTCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC	617
Qy	1816	TTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCTGGTTTATGGGCATCGCTGCAGC	1875
Db	618	TTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCTGGTTTATGGGCATCGCTGCAGC	677
Qy	1876	CAGACATGCCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGT	1935
Db	678	CAGACATGCCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGT	737
Qy	1936	GACTGCTTGCCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT	1995
Db	738	GACTGCTTGCCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT	797
Qy	1996	GGGAAAAAAGTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTTGAC	2055
Db	798	GGGAAAAAAGTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTTGAC	857
Qy	2056	AGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTGTCCCAGTGGCAGATTT	2115
Db	858	AGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGAAGTGTGTCCCAGTGGCAGATTT	917
Qy	2116	GCCCACTGGGGCCCAAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC	2175
Db	918	GCCCACTGGGGCCCAAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC	977
Qy	2176	GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA	2235
Db	978	GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA	1037
Qy	2236	TGTCCTCTAGGGTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCT	2295
Db	1038	TGTCCTCTAGGGTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCT	1097
Qy	2296	GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGCACTGT	2355

Db 1098 GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTCATGGGACGGCACTGT 1157
 Qy 2356 GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGTCTG 2415
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1158 GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGTCTG 1217
 Qy 2416 AACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA 2475
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1218 AACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA 1277
 Qy 2476 GCGAGATGTGATCAAGCTGGTGTTCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC 2535
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1278 GCGAGATGTGATCAAGCTGGTGTTCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC 1337
 Qy 2536 AGTACTGCTCTCCCTGCTGATTCCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTT 2595
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1338 AGTACTGCTCTCCCTGCTGATTCCTACCAAATCGGGGCCATTGCAGGCATCATCATTCTT 1397
 Qy 2596 GTCCTAGTTGTTCTCTTCCTACTGGCATTGTTTCATTATTTATAGACACA 2644
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1398 GTCCTAGTTGTTCTCTTCCTACTGGCATTGTTTCATTATTTATAGACACA 1446

RESULT 7

ABZ36212

ID ABZ36212 standard; cDNA; 1273 BP.

XX

AC ABZ36212;

XX

DT 10-FEB-2003 (first entry)

XX

DE Human secretory polynucleotide SPTM SEQ ID NO 376.

XX

KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
 KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
 KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
 KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
 KW anti-inflammatory; immunosuppressive; neuroprotective; nootropic;
 KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
 KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
 KW secretory polynucleotide; secretory protein; gene; ss.

XX

OS Homo sapiens.

XX

PN WO200283876-A2.

XX

PD 24-OCT-2002.

XX

PF 27-MAR-2002; 2002WO-US009921.

XX

PR 29-MAR-2001; 2001US-0280067P.

PR 29-MAR-2001; 2001US-0280068P.

PR 16-MAY-2001; 2001US-0291280P.

PR 17-MAY-2001; 2001US-0291829P.

PR 17-MAY-2001; 2001US-0291849P.

PR 19-JUN-2001; 2001US-0299428P.

PR 20-JUN-2001; 2001US-0299776P.

PR 20-JUN-2001; 2001US-0300001P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX

DR WPI; 2003-075543/07.

DR P-PSDB; ABP75770.

XX

PT New human secretory proteins and polynucleotides, useful for diagnosing,
PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
PT neurological disorders (e.g. Alzheimer's), or cell proliferations or
PT cancers.

XX

PS Claim 1; SEQ ID NO 376; 458pp + Sequence Listing; English.

XX

CC The invention relates to a secretory polynucleotide (designated sptm)
CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
CC naturally occurring polynucleotide sequence at least 90 % identical to
CC the polynucleotide sequence, a polynucleotide complementary to them or an
CC RNA equivalent of them. The polypeptide or polynucleotide are useful for
CC treating, preventing or diagnosing a disease or condition associated with
CC the expression of functional SPTM. These are particularly useful for
CC diagnosing, treating or preventing autoimmune/inflammatory disorders
CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,
CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
CC schizophrenia or amnesia), or cell proliferative disorders (e.g.
CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,
CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
CC breast, cervix or prostate). Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 1273 BP; 356 A; 311 C; 297 G; 309 T; 0 U; 0 Other;

Query Match 30.7%; Score 1049.4; DB 7; Length 1273;

Best Local Similarity 99.9%; Pred. No. 3.9e-294;

Matches 1050; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2373 AGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGTCTGAACAACTCCACCTGCGA 2432

|||||

Db 1 AGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGTCTGAACGACTCCACCTGCGA 60

Qy 2433 CCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGAGCGAGATGTGATCAAGC 2492

|||||

Db 61 CCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGAGCGAGATGTGATCAAGC 120

Qy 2493 TGGTGTTCATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACCACTGCTCTCCCTGC 2552

|||||

Db 121 TGGTGTTCATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACCACTGCTCTCCCTGC 180

Qy	2553	TGATTCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTTGTCTAGTTGTTCTCTT	2612
Db	181	TGATTCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTTGTCTAGTTGTTCTCTT	240
Qy	2613	CCTACTGGCATTGTTTCATTATTTATAGACACAAGCAGAAGGGAAAGGAATCAAGCATGCC	2672
Db	241	CCTACTGGCATTGTTTCATTATTTATAGACACAAGCAGAAGGGAAAGGAATCAAGCATGCC	300
Qy	2673	AGCAGTTACCTACACCCCTGCTATGAGGGTCGTCAATGCAGATTATACCATTTTCAGGAAC	2732
Db	301	AGCAGTTACCTACACCCCTGCTATGAGGGTCGTCAATGCAGATTATACCATTTTCAGGAAC	360
Qy	2733	CCTTCCTCACAGCAATGGTGGAAACGCTAATAGCCACTACTTCACCAATCCCAGTTACCA	2792
Db	361	CCTTCCTCACAGCAATGGTGGAAACGCTAATAGCCACTACTTCACCAATCCCAGTTACCA	420
Qy	2793	CACGCTCACCCAGTGTGCCACATCCCCTCACGTCAACAACAGGGACAGGATGACTGTCAC	2852
Db	421	CACGCTCACCCAGTGTGCCACATCCCCTCACGTCAACAACAGGGACAGGATGACTGTCAC	480
Qy	2853	GAAGTCAAAAAACAATCAACTGTTTGTGAATCTTAAAAATGTGAACCCTGGGAAGAGAGG	2912
Db	481	GAAGTCAAAAAACAATCAACTGTTTGTGAATCTTAAAAATGTGAACCCTGGGAAGAGAGG	540
Qy	2913	CCCTGTGGGGGACTGCACCTGGGACATTGCCGGCTGACTGGAAACATGGCGGCTACCTCAA	2972
Db	541	CCCTGTGGGGGACTGCACCTGGGACATTGCCGGCTGACTGGAAACATGGCGGCTACCTCAA	600
Qy	2973	CGAGCTCGGTGCTTTTGGACTTGACAGAAGCTATATGGGAAAATCCTTAAAAGACCTGGG	3032
Db	601	CGAGCTCGGTGCTTTTGGACTTGACAGAAGCTATATGGGAAAATCCTTAAAAGACCTGGG	660
Qy	3033	AAAGAATTCTGAATATAATTCAAGTAACTGCTCCCTAAGCAGTTCTGAGAACCCATATGC	3092
Db	661	AAAGAATTCTGAATATAATTCAAGTAACTGCTCCCTAAGCAGTTCTGAGAACCCATATGC	720
Qy	3093	CACTATTAAAGACCCACCTGTACTTATCCCGAAAAGCTCAGAGTGTGGTTATGTGGAGAT	3152
Db	721	CACTATTAAAGACCCACCTGTACTTATCCCGAAAAGCTCAGAGTGTGGTTATGTGGAGAT	780
Qy	3153	GAAATCGCCGGCACGAAGAGATTCCTCATATGCAGAGATCAATAACTCAACTTCAGCCAA	3212
Db	781	GAAATCGCCGGCACGAAGAGATTCCTCATATGCAGAGATCAATAACTCAACTTCAGCCAA	840
Qy	3213	CAGGAATGTCTATGAAGTTGAACCTACAGTGAGTGTGTCCAAGGAGTATTCAGCAATAA	3272
Db	841	CAGGAATGTCTATGAAGTTGAACCTACAGTGAGTGTGTCCAAGGAGTATTCAGCAATAA	900
Qy	3273	TGGGCGTCTCTCCAGGATCCATATGACCTCCCAAAGAACAGTCACATCCCTTGTCAATTA	3332
Db	901	TGGGCGTCTCTCCAGGATCCATATGACCTCCCAAAGAACAGTCACATCCCTTGTCAATTA	960
Qy	3333	TGACCTGCTGCCAGTCCGAGACAGTTCATCCTCCCCTAAGCAAGAGGACAGTGGAGGTAG	3392
Db	961	TGACCTGCTGCCAGTCCGAGACAGTTCATCCTCCCCTAAGCAAGAGGACAGTGGAGGTAG	1020
Qy	3393	CAGCAGCAACAGCAGCAGCAGCAGTGAATGA	3423

|||||
Db 1021 CAGCAGCAACAGCAGCAGCAGCAGTGAATGA 1051

RESULT 8

AAS91826

ID AAS91826 standard; cDNA; 1074 BP.

XX

AC AAS91826;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #27630.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

DR P-PSDB; ABG27639.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX

PS Claim 1; SEQ ID NO 27630; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 1074 BP; 241 A; 286 C; 303 G; 244 T; 0 U; 0 Other;

Query Match 18.1%; Score 618.4; DB 5; Length 1074;
Best Local Similarity 90.0%; Pred. No. 9.5e-169;
Matches 718; Conservative 0; Mismatches 1; Indels 79; Gaps 2;

```
Qy      412 GCCTGCGATGGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGG 471
          |||
Db      135 GCCTGCGATGGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGG 194

Qy      472 GCTCTGTGCAACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGC 531
          |||
Db      195 GCTCTGTGCAACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGC 254

Qy      532 TGCGAGGACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGC 591
          |||
Db      255 TGCGAGGACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGC 314

Qy      592 CAGAATGGAGCCACCTGCGACCACGTACGCGGGGAATGCCGCTGCCACCAGGATACACC 651
          |||
Db      315 CAGAATGGAGCCACCTGCGACCACGTACGCGGGGAATGCCGCTGCCACCAGGATACACC 374

Qy      652 GGAGCCTTCTGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGA 711
          |||
Db      375 GGAGCCTTCTGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGA 434

Qy      712 TGCCCTTGTCAAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCT 771
          |||
Db      435 TGCCCTTGTCAAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCT 494

Qy      772 GGCTGGAT----- 779
          |||
Db      495 GGCTGGATGTTGTCTTTCCCTGGCTGGAGGCCCATCTAATTTTCCAAGTCTCTTTGAATG 554

Qy      780 --GGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAA 837
          |||
Db      555 CAGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAA 614

Qy      838 GAATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGT 897
          |||
Db      615 GAATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGT 674

Qy      898 CCAGGATACACAGGGGAAC-----GGTGCCAGGATGAGTG 932
          |||
Db      675 CCAGGATACACAGGGGAACGAGCAGCAGTTCCGGATGTTAGAAAGGTGCCAGGATGAGTG 734

Qy      933 TCCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAA 992
          |||
Db      735 TCCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAA 794
```


Qy 993 GTGTTACACGTGAGCGGCGCATGCCTCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGA 1052
 |||
 Db 795 GTGTTACACGTGAGCGGCGCATGCCTCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGA 854
 Qy 1053 AGCACGCCTGTGTCCTGAGGGGCTCTACGGCATCAAATGTGACAAACGGTGTCCCTGCCA 1112
 |||
 Db 855 AGCACGCCTGTGTCCTGAGGGGCTCTACGGCATCAAATGTGACAAACGGTGTCCCTGCCA 914
 Qy 1113 CTTGAAAACACTCATAG 1130
 | |||
 Db 915 CTTGAAAACACTCATAG 932

RESULT 9

AAF27791

ID AAF27791 standard; cDNA; 3567 BP.

XX

AC AAF27791;

XX

DT 05-APR-2001 (first entry)

XX

DE Rat TANGO 272 coding sequence SEQ ID NO: 19.

XX

KW Membrane associated protein; secreted protein; human; mouse; rat;

KW INTERCEPT 340; MANGO 003; MANGO 347; TANGO 272; TANGO 295; TANGO 354;

KW TANGO 378; skeletal disorder; cardiovascular disorder; renal disorder;

KW haematopoietic disorder; neural disorder; hepatic disorder;

KW neoplastic disease; ss.

XX

OS Rattus sp.

XX

PN WO200100673-A1.

XX

PD 04-JAN-2001.

XX

PF 29-JUN-2000; 2000WO-US018198.

XX

PR 30-JUN-1999; 99US-00345464.

XX

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Barnes TM, Fraser CC, Wrighton N, Myers P, Busfield SJ, Sharp JD;

XX

DR WPI; 2001-050128/06.

DR P-PSDB; AAB66269.

XX

PT Isolated secreted or transmembrane proteins are used for diagnosis and

PT treatment of neoplastic and hematopoietic disorders e.g. T cell

PT disorders, cancer and tumors.

XX

PS Claim 1; Page 235-238; 294pp; English.

XX

CC The present invention provides the protein and coding sequences for a

CC number of membrane associated and secreted proteins from human, mouse and

CC rat. The proteins are designated INTERCEPT 340, MANGO 003, MANGO 347,

CC TANGO 272, TANGO 295, TANGO 254 and TANGO 378. The proteins are all

CC involved in signal transduction and the sequences can be used in the

CC treatment of cardiovascular, renal, hepatic, neural, neoplastic, skeletal
CC and haematopoietic disorders
XX
SQ Sequence 3567 BP; 690 A; 1115 C; 1000 G; 762 T; 0 U; 0 Other;

Query Match 17.8%; Score 608.4; DB 4; Length 3567;
Best Local Similarity 56.7%; Pred. No. 1.6e-165;
Matches 1226; Conservative 0; Mismatches 901; Indels 37; Gaps 4;

```
Qy      67 GCATCACCTCTGAATCTTGAAGACCCCTAATGTGTGTAGCCACTGGGAAAGCTACTCAGTG 126
      ||  | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      236 GCTGGAACACTCAACTCCAATGATCCCAATGTCTGTACCTTCTGGGAAAGCTTACCACG 295

Qy      127 ACTGTGCAAGAGTCATACCCACATCCCTTT---GATCAAATTTACTACACGAGCTGCACT 183
      ||  | | | | | | | | | | | | | | | | | | | | | | | |
Db      296 ACCACTAAGGAGTCCCACCTTCGCCCCCTTCAGCCTGCCCCAGCCGAGTCCTGCGACAGG 355

Qy      184 GACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTATCGA 243
      |  | | | | | | | | | | | | | | | | | | | | | | | |
Db      356 CCCTGGGAAGACCCCCACACCTGCGCTCAGCCTACGGTTGTCTACCGGACTGTGTACCGT 415

Qy      244 CATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTGTCTGGATTTTATGAAAGC 303
      || | | | | | | | | | | | | | | | | | | | | | | | |
Db      416 CAGGTGGTGAAGATGGACTCCCGCCACGCCTGCAGTGCTGTGGGGGTACTACGAGAGC 475

Qy      304 GGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTGCCTGTATTGCTCCA 363
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      476 AGTGGAGCCTGTGTCCCACTCTGTGCCAGGAGTGTGTCCACGGTGCCTGTGTGGCTCCT 535

Qy      364 AACACCTGTCTAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGATGGT 423
      ||  | | | | | | | | | | | | | | | | | | | | | | | |
Db      536 AATCGGTGCCAGTGTGCACCAGGCTGGCGGGGTGACGACTGTTCCAGTGAGTGTGCTCCT 595

Qy      424 GATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGCAAC 483
      |  | | | | | | | | | | | | | | | | | | | | | | | |
Db      596 GGAATGTGGGGACCACAGTGTGACAGGCTCTGCCTCTGTGGCAACAGCAGTTCCTGTGAT 655

Qy      484 CCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGC 543
      || | | | | | | | | | | | | | | | | | | | | | | | |
Db      656 CCCAGGAGTGGGGTGTGTTTTTGCCCTCTGGCCTGCAGCCCCCGACTGCCTTCAGCCT 715

Qy      544 TGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGAGCC 603
      ||  | | | | | | | | | | | | | | | | | | | | | | | |
Db      716 TGCCCCGATGGCCACTATGGTCTCTGCCAGTTTGATTGCCATTGC---TATGGGGCA 772

Qy      604 ACCTGCGACCACGTACGCGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTCTGT 663
      || | | | | | | | | | | | | | | | | | | | | | | | |
Db      773 TCCTGTGACCCCCGGGATGGAGCCTGCTTCTGCCCCCAGGGAGAACAGGACCCAGGGCA 832

Qy      664 GAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAA 723
      |  | | | | | | | | | | | | | | | | | | | | | | | |
Db      833 CTGATGGCTTCTTCTGCCCCAGAAC-----TTATCCTTGCCAA 870

Qy      724 AATGGAGGAGTGTGTATCATCAGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGC 783
      || | | | | | | | | | | | | | | | | | | | | | | | |
Db      871 AATGGAGGTGTTCTCAGGGCTCTCAAGGCTCCTGCAGCTGCCCACCGGGCTGGATGGGT 930
```

Qy	784	ACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGC	843
Db	931	GTCATCTGTTCCCTGCCATGCCAGAGGGTTTCCACGGACCCAACGTACTCAGGAATGT	990
Qy	844	CAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGA	903
Db	991	CGTTGCCACAATGGTGGCCTTTGTGACAGGTTTACTGGGCAGTGCCACTGTGCTCCTGGC	1050
Qy	904	TACACAGGGGAACGGTGCCAGGATGAGTGTCTCTGTTGGGACCTATGGCGTTCTCTGTGCT	963
Db	1051	TATATCGGGGATCGGTGCCGTGAAGAGTGCCCTGTGGGCCGCTTCGGTCAAGACTGTGCT	1110
Qy	964	GAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTTACCACGTGAGCGCGCATGCCTCTGT	1023
Db	1111	GAGACCTGTGACTGTGCTCCTGGCGCTCGTTGCTTTCTGCCAATGGCGCGTGTCTGTGC	1170
Qy	1024	GAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTCTGAGGGGCTCTACGGC	1083
Db	1171	GAACATGGCTTCACAGGCGACCGTGCCTGAGCGACTCTGTCCAGATGGCCGCTATGGT	1230
Qy	1084	ATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTCACCCCATG	1143
Db	1231	CTGAGCTGCCAAGATCCCTGCACCTGCGACCCAGAACACAGTCTCAGCTGCCACCCAATG	1290
Qy	1144	TCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCT	1203
Db	1291	CACGGCGAGTGCTCCTGCCAGCCAGGTTGGGCGGGCCTCCACTGCAACGAGAGCTGCCCT	1350
Qy	1204	CCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGT	1263
Db	1351	CAGGACACGCACGGAGCCGGTTGCCAGGAGCACTGCCTCTGTCTGCACGGCGGTGTTTGC	1410
Qy	1264	GACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACC	1323
Db	1411	CTCGCCGACAGCGGCCTCTGCCGGTGTGCACCTGGCTACACGGGACCTCACTGCGCTAAT	1470
Qy	1324	CCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAATGAT	1383
Db	1471	CTTTGTCCACCTAACACTTATGGGATCAACTGTTCTCTCCACTGCTCCTGTGAAAATGCC	1530
Qy	1384	GCAGTCTGCTCTCCTGTGGACGGGTCTTGACTTGCAAGGCAGGCTGGCACGGGGTGGAC	1443
Db	1531	ATTGCCTGCTCTCCTGTGACGGCACGTGCATCTGCAAGGAAGGTTGGCAGCGTGGTAAC	1590
Qy	1444	TGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGC	1503
Db	1591	TGCTCTGTGCCCTGTCCCCCTGGCACCTGGGGCTTCAGTTGCAATGCCAGTTGCCAGTGT	1650
Qy	1504	CTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGC	1563
Db	1651	GCCCACGAGGGAGTCTGCAGCCCCAACTGGAGCCTGTACTTGACCCCTGGGTGGCGT	1710
Qy	1564	GGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGC	1623
Db	1711	GGGGTTCACTGCCAACTTCCGTGCCCGAAGGGACAGTTTGGTGAAGGTTGTGCCAGTGTG	1770
Qy	1624	TGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCG	1683

KW Parkinson's disease; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 243. .3227
 FT /*tag= a
 FT /product= "CADHP-6"
 XX
 PN WO200259312-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 18-DEC-2001; 2001WO-US049206.
 XX
 PR 18-DEC-2000; 2000US-0256542P.
 PR 22-DEC-2000; 2000US-0259604P.
 PR 05-JAN-2001; 2001US-0260101P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Duggan BM, Xu Y, Lee EA, Lee S, Lu DAM, Warren BA, Yue H;
 PI Gietzen KJ, Honchell CD, Burford N, Baughn MR, Tang TY, Hillman JL;
 PI Gandhi AR, Kallick DA, Bandman O, Graul RC, Walia NK, Lu Y;
 PI Ramkumar J, Yao MG, Lal PG;
 XX
 DR WPI; 2002-590826/63.
 DR P-PSDB; AAG79417.
 XX
 PT New human cell adhesion proteins (CADHP) useful for treating, diagnosing
 PT and preventing diseases or conditions associated with the aberrant CADPH
 PT expression e.g. cancer, acquired immunodeficiency syndrome, Alzheimer's
 PT disease and epilepsy.
 XX
 PS Claim 5; Page 141-42; 149pp; English.
 XX
 CC The sequences given in ABA00054-63 encode novel human cell adhesion
 CC proteins (CADHP). The CADHP polypeptides and polynucleotides are useful
 CC in treating, diagnosing and preventing diseases or conditions associated
 CC with the decreased expression or overexpression of CADHP, e.g. immune
 CC system (acquired immunodeficiency syndrome, thymic dysplasia),
 CC neurological (Alzheimer's disease, Parkinson's disease, epilepsy),
 CC developmental (renal tubular acidosis, congenital glaucoma) and cell
 CC proliferative (cancer, atherosclerosis) disorders. They are also useful
 CC in assessing the effects of exogenous compounds on the expression of
 CC nucleic acid and amino acid sequences of CADHP. The CADHP or its
 CC fragments are useful in screening compounds for effectiveness as agonist
 CC or antagonist of the polypeptides, or in altering the expression of the
 CC target polynucleotide and compounds that specifically bind to or modulate
 CC the activity of the polypeptide. The protein encoded by this cDNA
 CC sequence shows homology to rat MEGF6
 XX
 SQ Sequence 3574 BP; 626 A; 1218 C; 1045 G; 685 T; 0 U; 0 Other;

Query Match 17.5%; Score 600.6; DB 6; Length 3574;
 Best Local Similarity 57.4%; Pred. No. 2.9e-163;
 Matches 1165; Conservative 0; Mismatches 849; Indels 15; Gaps 4;

Qy 62 GGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACT 121
 || || | ||| || | || ||| || ||| ||||| |||
 Db 289 GGCTGGCTGGAACCTCTCAACCCAGTGATCCCAATACCTGCAGCTTCTGGGAAAGCTTCA 348

Qy 122 CAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC- 180
 | || | |||| || | |||| | | || | |||
 Db 349 CTACCACCACCAAGGAGTCCCACTCCCGCCCTTCAGCCTGCTCCCTCAGAGCCCTGCG 408

Qy 181 --ACTGACATTCTAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCT 238
 | | || || || || || || || || || || || ||
 Db 409 AGCGGCCCTGGGAGGGCCCCATACCTGCCCCAGCCACGGTTGTATACGGACCGTGT 468

Qy 239 ATCGACATGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCTGGATTTTATG 298
 | || || | || |||| | | | |||| || | || | ||||
 Db 469 ACCGTCAGGTGGTGAAGACGGACCACGCCAGCGCTGCAGTGCTGCCATGGCTTCTATG 528

Qy 299 AAAGCGGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTG 358
 | ||| ||| | ||||| ||||| | | ||||| ||| ||| | |
 Db 529 AGAGCAGGGGGTTCTGTGTCCCGCTCTGTGCCAGGAGTGTGTCCATGGCCGTTGTGTGG 588

Qy 359 CTCCAAACACCTGTGAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCG 418
 | || || | || || |||| || |||| | || | |||| ||||| || |
 Db 589 CACCAATCAGTGCCAATGTGTGCCAGGCTGGCGGGGCGACGACTGTTCCAGTGAGTGTG 648

Qy 419 ATGGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGT 478
 | |||| || || || || || | || || || || || ||
 Db 649 CCCCAGGAATGTGGGGGCCACAGTGTGACAAGCCCTGCAGCTGCGGCAACAACAGCTCGT 708

Qy 479 GCAACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGG 538
 | | |||| | |||| || || || || || || || || || ||
 Db 709 GTGATCCCAAGAGTGGGGTATGTTCTTGCCCTTCTGGTCTGCAGCCCCGAAGTGCCTTC 768

Qy 539 ACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATG 598
 | | |||| | || |||| | || || | ||||| ||| |||
 Db 769 AGCCCTGTACCCCTGGCTACTATGGCCCTGCCTGCCAGTTCGCTGCCAGTGCC--ATG 825

Qy 599 GAGCCACCTGCGACCACGTACGCGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCT 658
 | || ||||| || | || || | || |||| || | || || ||
 Db 826 GGGCACCTGCGATCCCAGACTGGAGCCTGCTTCTGCCCCGAGAGAGAACTGGGCCCA 885

Qy 659 TCTGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTT 718
 ||||| | |||| | || | || || || || || || || ||
 Db 886 GCTGTGACGTGTCCTGTTCCAGGGCACTTCTGGCTTCTTCTGCCCCAGCACCCATCCTT 945

Qy 719 GTCAAAATGGAGGAGTGTGTATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGA 778
 | ||||| || || || | || || || || || || || || || ||
 Db 946 GCCAAAATGGAGGTGTCTTCAAACCCACAGGGCTCCTGCAGCTGCCCCCTGGCTGGA 1005

Qy 779 TGGGCACAGTGTGTGGTCAGCCTTGCCCCAGGGTGCCTTTGGAAAGAACTGTTCCCAAG 838
 ||||| | || | || |||| |||| || || || || || || ||
 Db 1006 TGGGCACCATCTGCTCCCTGCCCTGCCAGAGGGCTTTCACGGACCAACTGCTCCCAGG 1065

Qy 839 AATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTC 898
 |||| | |||| || || || || || || || || || || || ||
 Db 1066 AATGTCGCTGCCACAACGGCGGCCTCTGTGACCGATTCACTGGGCAGTGCCGCTGCGCTC 1125

Qy	899	CAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCT	958
Db	1126	CGGGTTACACTGGGGATCGGTGCCGGGAGGAGTGCCCGGTGGGCCGCTTTGGGCAGGACT	1185
Qy	959	GTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCC	1018
Db	1186	GTGCTGAGACGTGCGACTGCGCCCCGGACGCCCCTTGCTTCCCGGCCAACGGCGCATGTC	1245
Qy	1019	TCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTTGAGGGGCTCT	1078
Db	1246	TGTGCGAACACGGCTTCACTGGGGACCGCTGCACGGATCGCCTCTGCCCCGACGGCTTCT	1305
Qy	1079	ACGGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTCACC	1138
Db	1306	ACGGTCTCAGCTGCCAGGCCCTTGCACCTGCGACCGGGAGCACAGCCTCAGCTGCCACC	1365
Qy	1139	CCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACAT	1198
Db	1366	CGATGAACGGGGAGTGTCTCTGCCTGCCGGGCTGGGCGGGCCTCCACTGCAACGAGAGCT	1425
Qy	1199	GTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAG	1258
Db	1426	GCCCGCAGGACACGCATGGGCCAGGGTGCCAGGAGCACTGTCTCTGCCTGCACGGTGGCG	1485
Qy	1259	ACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCT	1318
Db	1486	TCTGCCAGGCTACCAGCGGCCTCTGTCACTGCGCGCCGGGTACACGGGGCCCTCACTGTG	1545
Qy	1319	CTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAA	1378
Db	1546	CTAGTCTTTGTCTCCTGACACCTACGGTGTCAACTGTTCTGCACGCTGCTCATGTGAAA	1605
Qy	1379	ATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGG	1438
Db	1606	ATGCCATCGCCTGCTCACCCATCGACGGCGAGTGCCTCTGCAAGGAAGGTTGGCAGCGTG	1665
Qy	1439	TGGACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCC	1498
Db	1666	GTAAGTGTCTGTGCCCTGCCACCCGGAACCTGGGGCTTCAGTTGCAATGCCAGCTGCC	1725
Qy	1499	AGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGAT	1558
Db	1726	AGTGTGCCCATGAGGCAGTCTGCAGCCCCCAAACCTGGAGCCTGTACCTGCACCCCTGGGT	1785
Qy	1559	GGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTG	1618
Db	1786	GGCATGGGGCCCACTGCCAGCTGCCCTGTCCGAAGGGGCAGTTTGGAGAAGGTTGTGCCA	1845
Qy	1619	AGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCC	1678
Db	1846	GTCGCTGTGACTGTGACCACTCTGATGGCTGTGACCCTGTTTCATGGACGCTGTCAGTGCC	1905
Qy	1679	TCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCA	1738
Db	1906	AGGCTGGCTGGATGGGTGCCCGCTGCCACCTGTCTGCCCCTGAGGGCTTATGGGGAGTCA	1965
Qy	1739	ACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCT	1798

```

      |||| | |||| ||| || ||||| | || ||||| ||||| ||
Db      1966 ACTGTAGCAACACCTGCACCTGCAAGAATGGGGGCACCTGTCTCCCTGAGAATGGCAACT 2025

Qy      1799 GCGAGTGTGCACCAGGCTTCCGAGGCACCACTTGTCTCAGAGGATCTGCTCCCCTGGTTTTT 1858
      ||| ||||| || ||||| ||| || | || ||||| ||| ||||| |
Db      2026 GCGTGTGTGCACCCGGATTCCGGGGCCCCCTCTGCCAGAGATCCTGTCTCAGCCTGGCCGCT 2085

Qy      1859 ATGGGCATCGCTGCAGCCAGACATGCCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACC 1918
      |||| | |||| | | ||||| || || | | |||||
Db      2086 ATGGCAAACGCTGTGTGC-----CCTGCAAGTGCCTAACCCTCC---TTCTGCCACC 2136

Qy      1919 ACATCACCGGCCTGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGT 1978
      | | ||| || ||||| || ||||| ||||| || |||| | |
Db      2137 CCTCGAACGGGACCTGCTACTGCCTGGCTGGCTGGACAGGCCCCGACTGCTCCCAGCGCT 2196

Qy      1979 GTCCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAA 2038
      | || || | ||||| ||||| | || || || ||
Db      2197 GCCCTCTGGGGACATTTGGTGCTAACTGCTCCCAGCCATGCCAGTGTGGTCCTGGAGAAA 2256

Qy      2039 CCTGTAACCCATTGACAGATCTTGTCTAGTGTACCCCGGTTGGATTGG 2087
      || |||| | | ||| ||| ||| || ||||
Db      2257 AGTGCCACCCAGAGACTGGGGCCTGTGTATGTCCCCCAGGGCACAGTGG 2305

```

RESULT 11

AAS86746

ID AAS86746 standard; cDNA; 1402 BP.

XX

AC AAS86746;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #22550.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

DR P-PSDB; ABG22559.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX

PS Claim 1; SEQ ID NO 22550; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 1402 BP; 356 A; 335 C; 413 G; 298 T; 0 U; 0 Other;

Query Match 17.0%; Score 581.6; DB 5; Length 1402;
Best Local Similarity 88.5%; Pred. No. 5.7e-158;
Matches 710; Conservative 0; Mismatches 9; Indels 83; Gaps 4;

```
Qy      412 GCCTGCGATGGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGG 471
      |||
Db      135 GCCTGCGATGGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGG 194

Qy      472 GCTCTGTGCAA-CCCCATCACCGGGGCTTGC---CACTGTGCTGCGGGCTTCCGGGGCTG 527
      |||
Db      195 GCTCTGTGCAACCCCCATCACCGGGGCTTGCCCACTGTGGCTGCGGGGCTTCCGGGGCTG 254

Qy      528 GCGCTGCGAGGACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCA 587
      |||
Db      255 GCGCTGCGAGGACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCA 314

Qy      588 GTGCCAGAATGGAGCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCACCAGGATA 647
      |||
Db      315 GTGCCAGAATGGAGCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCACCAGGATA 374

Qy      648 CACCGGAGCCTTCTGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCA 707
      |||
Db      375 CACCGGAGCCTTCTGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCA 434

Qy      708 GAGATGCCCTTGTCAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCC 767
      |||
Db      435 GAGATGCCCTTGTCAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCC 494
```

Qy	768	TTCTGGCTGGAT-----	779
Db	495	TTCTGGCTGGATGTTGTCTTCCCTGGCTGGAGGCCCATCTAATTTTCCAAGTCTCTTTG	554
Qy	780	-----GGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTC	833
Db	555	AATGCAGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTC	614
Qy	834	CCAAGAATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTG	893
Db	615	CCAAGAATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTG	674
Qy	894	CAGTCCAGGATACACAGGGGAAC-----GGTGCCAGGATG	928
Db	675	CAGTCCAGGATACACAGGGGAACGAGCAGCAGTTCCGGATGTTAGAAAGGTGCCAGGATG	734
Qy	929	AGTGTCTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAG	988
Db	735	AGTGTCTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAG	794
Qy	989	GGAAGTGTTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAGGCTTTGCTGGCGAGCGCT	1048
Db	795	GGAAGTGTTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAGGCTTTGCTGGCGAGCGCT	854
Qy	1049	GCGAAGCACGCCTGTGTCTCTGAGGGGCTCTACGGCATCAAATGTGACAAACGGTGTCCCT	1108
Db	855	GCGAAGCACGCCTGTGTCTCTGAGGGGCTCTACGGCATCAAATGTGACAAACGGTGTCCCT	914
Qy	1109	GCCACTTGAAAAACACTCATAG	1130
Db	915	GCCACCTTGAAAAACACTCATAG	936

RESULT 12

AAS72218

ID AAS72218 standard; cDNA; 936 BP.

XX

AC AAS72218;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #8022.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG08031.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 8022; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 936 BP; 200 A; 239 C; 258 G; 239 T; 0 U; 0 Other;

Query Match 16.5%; Score 563.8; DB 5; Length 936;
 Best Local Similarity 97.9%; Pred. No. 6.8e-153;
 Matches 571; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1130 GCTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTA 1189
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 95 GCTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTA 154
 Qy 1190 ATGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAA 1249
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 155 ATGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAA 214
 Qy 1250 ATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGACCTGTGCCCCAGGATTCAAAGGAA 1309
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 215 ATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGACCTGTGCCCCAGGATTCAAAGGAA 274
 Qy 1310 TTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTTCCTCTCGTGTG 1369
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 275 TTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTTCCTCTCGTGTG 334

QY 1370 GCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCT 1429
 |||
 Db 335 GCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCT 394
 QY 1430 GGCACGGGGTGGACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACT 1489
 |||
 Db 395 GGCACGGGGTGGACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACT 454
 QY 1490 TAACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTG 1549
 |||
 Db 455 TAACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTG 514
 QY 1550 CACCTGGATGGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGA 1609
 |||
 Db 515 CACCTGGATGGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGA 574
 QY 1610 ACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGGCCATT 1669
 |||
 Db 575 ACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGGCCATT 634
 QY 1670 GCCGCTGCCTCCCGGGATGGTCAGGTGTCCACTGTGACAGCGT 1712
 |||
 Db 635 GCCGCTGCCTCCCGGGATGGTCAGTGTTCACCGGAAATGGTGT 677

RESULT 13

ACD05889

ID ACD05889 standard; cDNA; 936 BP.

XX

AC ACD05889;

XX

DT 06-AUG-2003 (first entry)

XX

DE Novel human contig #63.

XX

KW Human; angiogenesis; cytokine; cell proliferation; pluripotent;
 KW cell differentiation; totipotent; stem cell; transplantation; bio-sensor;
 KW neuroepithelial cell; autoimmune disease; neural cell; genetic disorder;
 KW nerve; brain tissue; central nervous system disease;
 KW peripheral nervous system disease; neuropathy; haematopoiesis; bone;
 KW myeloid disorder; lymphoid cell disorder; platelet disorder; tendon;
 KW regeneration; cartilage; tendon; ligament; nerve tissue growth;
 KW tissue repair; wound healing; burn; ulcer; osteoporosis; cancer;
 KW osteoarthritis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung fibrosis; liver fibrosis; reperfusion injury;
 KW immune deficiency; infection; autoimmune disorder; allergic reaction;
 KW thrombolysis; thrombosis; coagulation disorder; hereditary disorder;
 KW biorhythm; circadian cycle; fertility; metabolism; catabolism; anabolism;
 KW nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
 KW haemostatic; vulnerary; antiulcer; osteopathic; antiarthritic;
 KW vasotropic; immunostimulant; antibacterial; fungicide; immunosuppressive;
 KW antirheumatic; antidiabetic; antiasthmatic; cytostatic; virucide;
 KW expressed sequence tag; EST; ss.

XX

OS Homo sapiens.

XX

PN WO2003023013-A2.
XX
PD 20-MAR-2003.
XX
PF 13-SEP-2002; 2002WO-US029001.
XX
PR 13-SEP-2001; 2001US-0322511P.
PR 12-SEP-2002; 2002US-00243552.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Yang Y, Wang Z, Weng G, Ma Y;
XX
DR WPI; 2003-313249/30.
DR P-PSDB; ABO00812.
XX
PT Novel nucleic acids and polypeptides for diagnosis, treatment of central
PT and peripheral nervous system diseases and neuropathies, such as
PT Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
PT lateral sclerosis.
XX
PS Example 2; SEQ ID NO 735; 300pp; English.
XX
CC The present invention relates to the isolation of novel human
CC polynucleotide sequences and their encoding polypeptides. The novel
CC polypeptides exhibit activities relating to angiogenesis, cytokine, cell
CC proliferation, cell differentiation, antiinflammatory, and stem cell
CC growth factor activities. The polypeptides are involved in the
CC proliferation, differentiation and survival of pluripotent and totipotent
CC stem cells, and are useful for re-engineering damaged or diseased
CC tissues, transplantation, manufacture of bio-pharmaceuticals and
CC development of bio-sensors. The polypeptides can be used to manipulate
CC stem cells in culture to give rise to neuroepithelial cells that can be
CC used to augment or replace cells damaged by illness, autoimmune disease,
CC accidental damage or genetic disorders. The polypeptides induce the
CC proliferation of neural cells and regeneration of nerve and brain tissue
CC and are useful for the treatment of central and peripheral nervous system
CC diseases and neuropathies, such as Alzheimer's, Parkinson's disease,
CC Huntington's disease, amyotrophic lateral sclerosis (ALS). The
CC polypeptides are also involved in chemotactic or chemokinetic activity,
CC regulation of haematopoiesis and are useful for treating myeloid or
CC lymphoid cell disorders, platelet disorders such as thrombocytopaenia and
CC for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, in tissue repair, healing of burns, incisions, ulcers, for
CC treating osteoporosis, osteoarthritis, bone degenerative disorders, and
CC periodontal disease. The polypeptides are also useful for gut protection
CC or regeneration and treatment of lung or liver fibrosis, reperfusion
CC injury in various tissues, various immune deficiencies and disorders
CC including severe combined immunodeficiency (SCID), bacterial or fungal
CC infections, autoimmune disorders (e.g. multiple sclerosis, rheumatoid
CC arthritis, diabetes mellitus, myasthenia gravis), allergic reactions and
CC conditions, such as asthma or other respiratory problems. The
CC polypeptides are involved in thrombolysis or thrombosis and are useful in
CC treatment of various coagulation disorders (including hereditary
CC disorders such as haemophilia) or to enhance coagulation and other
CC haemostatic events in treating wounds resulting from trauma, surgery or
CC other causes. The polypeptides exhibit immune stimulating or immune

CC suppressing activity, and are useful for treating autoimmune diseases or
CC cancer. They also inhibit the growth, infection or function of infectious
CC agents such as bacteria, fungi, viruses, effect biorhythms or circadian
CC cycles of rhythms, fertility of male or female subjects, metabolism,
CC catabolism, and anabolism. ACD05827-ACD06027 represent novel contigs
CC assembled using expressed sequence tag (EST) sequences as seeds. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 936 BP; 200 A; 239 C; 258 G; 239 T; 0 U; 0 Other;

Query Match 16.5%; Score 563.8; DB 7; Length 936;
Best Local Similarity 97.9%; Pred. No. 6.8e-153;
Matches 571; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy      1130 GCTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTA 1189
          |||
Db      95  GCTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTA 154

Qy      1190 ATGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAA 1249
          |||
Db     1155 ATGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAA 214

Qy      1250 ATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAA 1309
          |||
Db      215 ATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAA 274

Qy      1310 TTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTG 1369
          |||
Db      275 TTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTG 334

Qy      1370 GCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCT 1429
          |||
Db      335 GCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCT 394

Qy      1430 GGCACGGGGTGGACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACT 1489
          |||
Db      395 GGCACGGGGTGGACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACT 454

Qy      1490 TAACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTG 1549
          |||
Db      455 TAACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTG 514

Qy      1550 CACCTGGATGGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGA 1609
          |||
Db      515 CACCTGGATGGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGA 574

Qy      1610 ACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATT 1669
          |||
Db      575 ACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATT 634

Qy      1670 GCCGCTGCCTCCCGGGATGGTCAGGTGTCCACTGTGACAGCGT 1712
          |||
Db      635 GCCGCTGCCTCCCGGGATGGTCAGGTGTTCACCGGAAATGGTGT 677
```

RESULT 14

ACD05589

ID ACD05589 standard; cDNA; 936 BP.

XX

AC ACD05589;

XX

DT 06-AUG-2003 (first entry)

XX

DE cDNA encoding novel human polypeptide #99.

XX

KW Human; angiogenesis; cytokine; cell proliferation; pluripotent;
 KW cell differentiation; totipotent; stem cell; transplantation; bio-sensor;
 KW neuroepithelial cell; autoimmune disease; neural cell; genetic disorder;
 KW nerve; brain tissue; central nervous system disease;
 KW peripheral nervous system disease; neuropathy; haematopoiesis; bone;
 KW myeloid disorder; lymphoid cell disorder; platelet disorder; tendon;
 KW regeneration; cartilage; tendon; ligament; nerve tissue growth;
 KW tissue repair; wound healing; burn; ulcer; osteoporosis; cancer;
 KW osteoarthritis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung fibrosis; liver fibrosis; reperfusion injury;
 KW immune deficiency; infection; autoimmune disorder; allergic reaction;
 KW thrombolysis; thrombosis; coagulation disorder; hereditary disorder;
 KW biorhythm; circadian cycle; fertility; metabolism; catabolism; anabolism;
 KW nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
 KW haemostatic; vulnerary; antiulcer; osteopathic; antiarthritic;
 KW vasotropic; immunostimulant; antibacterial; fungicide; immunosuppressive;
 KW antirheumatic; antidiabetic; antiasthmatic; cytostatic; virucide; gene;
 KW ss.

XX

OS Homo sapiens.

XX

PN WO2003023013-A2.

XX

PD 20-MAR-2003.

XX

PF 13-SEP-2002; 2002WO-US029001.

XX

PR 13-SEP-2001; 2001US-0322511P.

PR 12-SEP-2002; 2002US-00243552.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Yang Y, Wang Z, Weng G, Ma Y;

XX

DR WPI; 2003-313249/30.

DR P-PSDB; ABO00512.

XX

PT Novel nucleic acids and polypeptides for diagnosis, treatment of central
 PT and peripheral nervous system diseases and neuropathies, such as
 PT Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 PT lateral sclerosis.

XX

PS Claim 1; SEQ ID NO 99; 300pp; English.

XX

CC The present invention relates to the isolation of novel human
 CC polynucleotide sequences and their encoding polypeptides. The novel
 CC polypeptides exhibit activities relating to angiogenesis, cytokine, cell

proliferation, cell differentiation, antiinflammatory, and stem cell growth factor activities. The polypeptides are involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells, and are useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and development of bio-sensors. The polypeptides can be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptides induce the proliferation of neural cells and regeneration of nerve and brain tissue and are useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis (ALS). The polypeptides are also involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and are useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopaenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, and periodontal disease. The polypeptides are also useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis), allergic reactions and conditions, such as asthma or other respiratory problems. The polypeptides are involved in thrombolysis or thrombosis and are useful in treatment of various coagulation disorders (including hereditary disorders such as haemophilia) or to enhance coagulation and other haemostatic events in treating wounds resulting from trauma, surgery or other causes. The polypeptides exhibit immune stimulating or immune suppressing activity, and are useful for treating autoimmune diseases or cancer. They also inhibit the growth, infection or function of infectious agents such as bacteria, fungi, viruses, effect biorhythms or circadian cycles of rhythms, fertility of male or female subjects, metabolism, catabolism, and anabolism. ACD05491-ACD05826 represent the novel cDNA sequences of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 936 BP; 200 A; 239 C; 258 G; 239 T; 0 U; 0 Other;

Query Match 16.5%; Score 563.8; DB 7; Length 936;
Best Local Similarity 97.9%; Pred. No. 6.8e-153;
Matches 571; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

[illegible]


```

Db      215 ATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAA 274
Qy      1310 TTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTG 1369
          |||
Db      275 TTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTG 334
Qy      1370 GCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCT 1429
          |||
Db      335 GCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCT 394
Qy      1430 GGCACGGGGTGGACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTA ACT 1489
          |||
Db      395 GGCACGGGGTGGACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTA ACT 454
Qy      1490 TAACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTG 1549
          |||
Db      455 TAACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTG 514
Qy      1550 CACCTGGATGGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGA 1609
          |||
Db      515 CACCTGGATGGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGA 574
Qy      1610 ACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATT 1669
          |||
Db      575 ACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATT 634
Qy      1670 GCCGCTGCCTCCCGGGATGGTCAGGTGTCCACTGTGACAGCGT 1712
          |||
Db      635 GCCGCTGCCTCCCGGGATGGTCAGTGTTACCGGAAATGGTGT 677

```

RESULT 15

AAS72219

ID AAS72219 standard; cDNA; 2295 BP.

XX

AC AAS72219;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #8023.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG08032.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 8023; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2295 BP; 610 A; 543 C; 602 G; 540 T; 0 U; 0 Other;

Query Match 16.2%; Score 554; DB 5; Length 2295;
 Best Local Similarity 97.3%; Pred. No. 8.2e-150;
 Matches 574; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

Qy 2834 GGGACAGGATGACTGTACGAAGTCAAAAAACAATCAACTGTTTGTGAATCTTAAAAATG 2893
 |||| | ||| | |||||
 Db 575 GGGAGGGACCAACTGCTCCAGTGTCAAAAAACAATCAACTGTTTGTGAATCTTAAAAATG 634

 Qy 2894 TGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTGCCGGCTGACTGGA 2953
 |||||
 Db 635 TGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTGCCGGCTGACTGGA 694

 Qy 2954 AACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGAAGCTATATGGGAA 3013
 |||||
 Db 695 AACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGAAGCTATATGGGAA 754

 Qy 3014 AATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAGTCTCCCTAAGCA 3073
 |||||
 Db 755 AATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAGTCTCCCTAAGCA 814

 Qy 3074 GTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATCCCGAAAAGCTCAG 3133

```

      |||
Db      815 GTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATCCCGAAAAGCTCAG 874
      |||
Qy      3134 AGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCATATGCAGAGATCA 3193
      |||
Db      875 AGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCATATGCAGAGATCA 934
      |||
Qy      3194 ATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACAGTGAGTGTGTCC 3253
      |||
Db      935 ATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACAGTGAGTGTGTCC 994
      |||
Qy      3254 AAGGAGTATTCAGCAATAATGGGCGTCTCTCCAGGATCCATATGACCTCCCAAAGAACA 3313
      |||
Db      995 AAGGAGTATTCAGCAATAATGGGCGTCTTTCCAGGATCCATATGACCTCCC-AAGAACA 1053
      |||
Qy      3314 GTCACATCCCTTGTCAATTATGACCTGCTGCCAGTCCGAGACAGTTCATCCTCCCTAAGC 3373
      |||
Db      1054 GTCACATCCCTTGTCAATTATGACCTGCTGCCAGTCCGAGACAGTTCATCCTCCCTTAAGC 1113
      |||
Qy      3374 AAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGCAGTGAATGA 3423
      |||
Db      1114 AAGAGGACAGTGGTGGTAGCAGCAGCAACAGCAGCAGCAGCAGTGAATGA 1163

```

Search completed: March 30, 2004, 02:37:09
Job time : 1291 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 02:09:56 ; Search time 235 Seconds
(without alignments)
8083.391 Million cell updates/sec

Title: US-10-092-390-1
Perfect score: 3423
Sequence: 1 atgggtattttctttgaactc.....gcagcagcagcagtgaatga 3423

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	8		Query		DB	ID	Description
	Score	Match	Length	Length			
c	1	1425.8	41.7	1448	4	US-09-130-491-9	Sequence 9, Appli
	2	524.6	15.3	5197	4	US-09-833-381-1910	Sequence 1910, Ap
	3	165.8	4.8	1578	4	US-09-130-491-10	Sequence 10, Appl
	4	156	4.6	1464	3	US-09-188-930-255	Sequence 255, App
	5	156	4.6	1464	4	US-09-312-283C-255	Sequence 255, App
	6	156	4.6	1635	4	US-09-312-283C-73	Sequence 73, Appl
	7	155.2	4.5	1633	3	US-09-188-930-73	Sequence 73, Appl
	8	134.4	3.9	393	4	US-09-833-381-1076	Sequence 1076, Ap
	9	66.2	1.9	553	4	US-09-621-976-1192	Sequence 1192, Ap
	10	55.2	1.6	4138	1	US-08-323-474-1	Sequence 1, Appli
	11	55.2	1.6	4138	5	PCT-US93-06093-1	Sequence 1, Appli

	12	54	1.6	3845	2	US-08-220-240A-4	Sequence 4, Appli
	13	54	1.6	4175	1	US-07-934-393B-1	Sequence 1, Appli
	14	54	1.6	4175	1	US-08-278-089A-1	Sequence 1, Appli
	15	54	1.6	4175	2	US-08-838-957A-1	Sequence 1, Appli
	16	54	1.6	4176	1	US-08-278-089A-5	Sequence 5, Appli
	17	54	1.6	4176	2	US-08-838-957A-5	Sequence 5, Appli
c	18	49	1.4	3489	2	US-08-728-323A-1	Sequence 1, Appli
c	19	49	1.4	3489	4	US-09-298-568-1	Sequence 1, Appli
c	20	49	1.4	3489	4	US-09-410-399-1	Sequence 1, Appli
	21	49	1.4	32207	2	US-08-770-379-20	Sequence 20, Appl
	22	49	1.4	32207	3	US-08-757-669A-20	Sequence 20, Appl
	23	49	1.4	32207	4	US-09-230-371A-20	Sequence 20, Appl
	24	48.4	1.4	1386	4	US-09-252-991A-11863	Sequence 11863, A
	25	48.4	1.4	2304	4	US-09-252-991A-11802	Sequence 11802, A
	26	47.4	1.4	2005	4	US-09-635-872A-18	Sequence 18, Appl
	27	47.4	1.4	2005	4	US-09-636-077A-18	Sequence 18, Appl
	28	47.4	1.4	2005	4	US-09-636-060C-18	Sequence 18, Appl
	29	47.4	1.4	2005	4	US-09-986-552-18	Sequence 18, Appl
c	30	46.4	1.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	31	46.4	1.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	32	45.6	1.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	33	44.4	1.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	34	40.6	1.2	2144	3	US-08-834-306-15	Sequence 15, Appl
	35	40.6	1.2	2144	3	US-08-993-674A-15	Sequence 15, Appl
	36	40.6	1.2	2144	4	US-09-256-976-15	Sequence 15, Appl
	37	40.2	1.2	13857	4	US-09-620-312D-75	Sequence 75, Appl
	38	39.8	1.2	1383	4	US-09-635-872A-22	Sequence 22, Appl
	39	39.8	1.2	1383	4	US-09-636-077A-22	Sequence 22, Appl
	40	39.8	1.2	1383	4	US-09-636-060C-22	Sequence 22, Appl
	41	39.8	1.2	1383	4	US-09-986-552-22	Sequence 22, Appl
	42	39.8	1.2	2183	4	US-09-635-872A-7	Sequence 7, Appli
	43	39.8	1.2	2183	4	US-09-636-077A-7	Sequence 7, Appli
	44	39.8	1.2	2183	4	US-09-636-060C-7	Sequence 7, Appli
	45	39.8	1.2	2183	4	US-09-986-552-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-130-491-9

; Sequence 9, Application US/09130491

; Patent No. 6416974

; GENERAL INFORMATION:

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Goodearl, Andrew D.J.

; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83

; FILE REFERENCE: 09404/041001

; CURRENT APPLICATION NUMBER: US/09/130,491

; CURRENT FILING DATE: 1998-08-07

; EARLIER APPLICATION NUMBER: US 60/058,108

; EARLIER FILING DATE: 1997-09-05

; EARLIER APPLICATION NUMBER: US 60/054,961

; EARLIER FILING DATE: 1997-08-06

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 9

; LENGTH: 1448
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-130-491-9

Query Match 41.7%; Score 1425.8; DB 4; Length 1448;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1427; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1216	GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT	1275
Db	18	GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT	77
Qy	1276	GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG	1335
Db	78	GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG	137
Qy	1336	GGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT	1395
Db	138	GGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT	197
Qy	1396	CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGA	1455
Db	198	CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGA	257
Qy	1456	TGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGA	1515
Db	258	TGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGA	317
Qy	1516	GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC	1575
Db	318	GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC	377
Qy	1576	GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC	1635
Db	378	GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC	437
Qy	1636	CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCCGGATGGTCAGGT	1695
Db	438	CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCCGGATGGTCAGGT	497
Qy	1696	GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC	1755
Db	498	GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC	557
Qy	1756	TACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC	1815
Db	558	TACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC	617
Qy	1816	TTCCGAGGCACCACTTGTCTAGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCTGCAGC	1875
Db	618	TTCCGAGGCACCACTTGTCTAGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCTGCAGC	677
Qy	1876	CAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATACCCGGCCTGTGT	1935
Db	678	CAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATACCCGGCCTGTGT	737

Qy	1936	GACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT	1995
Db	738	GACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT	797
Qy	1996	GGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTGAC	2055
Db	798	GGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTGAC	857
Qy	2056	AGATCTTGTCTAGTGTACCCCGGTTGGATTGGCAGTGACTGCTCTCAACCATGTCCACCT	2115
Db	858	AGATCTTGTCTAGTGTACCCCGGTTGGATTGGCAGTGACTGCTCTCAACCATGTCCACCT	917
Qy	2116	GCCCACTGGGGCCCAAACATGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC	2175
Db	918	GCCCACTGGGGCCCAAACATGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC	977
Qy	2176	GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA	2235
Db	978	GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA	1037
Qy	2236	TGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCT	2295
Db	1038	TGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCT	1097
Qy	2296	GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTCATGGGACGGCACTGT	2355
Db	1098	GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTCATGGGACGGCACTGT	1157
Qy	2356	GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCCAGATATGTGATTGTCTG	2415
Db	1158	GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCCAGATATGTGATTGTCTG	1217
Qy	2416	AACAACCTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA	2475
Db	1218	AACAACCTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA	1277
Qy	2476	GCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC	2535
Db	1278	GCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC	1337
Qy	2536	AGTACTGCTCTCCCTGCTGATTCCCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTT	2595
Db	1338	AGTACTGCTCTCCCTGCTGATTCCCTACCAAATCGGGGCCATTGCAGGCATCATCATTCTT	1397
Qy	2596	GTCCTAGTTGTTCTCTTCTACTGGCATTGTTTCATTATTTATAGACACA	2644
Db	1398	GTCCTAGTTGTTCTCTTCTACTGGCATTGTTTCATTATTTATAGACACA	1446

RESULT 2

US-09-833-381-1910/c

; Sequence 1910, Application US/09833381

; Patent No. 6672186

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

```
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1910
; LENGTH: 5197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(5197)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1910
```

```
Query Match          15.3%; Score 524.6; DB 4; Length 5197;
Best Local Similarity 58.0%; Pred. No. 1.3e-152;
Matches 972; Conservative 0; Mismatches 694; Indels 11; Gaps 2;
```

```
Qy      531 CTGCGAGGACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTG 590
      |||| | | |||| | ||| ||||| | ||| || | |||||
Db      4243 CTGCCTTCAGCCCTGTACCCCTGGCTACTATGGCCCTGCCTGCCAGTTCGCTGCCAGTG 4184

Qy      591 CCAGAATGGAGCCACCTGCGACCACGTCACGGGGGAATGCCGCTGCCCACCAGGATACAC 650
      || |||| || ||||| | | || | | ||| ||||| ||| ||
Db      4183 CC---ATGGGGCACCCCTGCGATCCCCAGACTGGAGCCTGCTTCTGCCCCGAGAGAGAAC 4127

Qy      651 CGGAGCCTTCTGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAG 710
      || || ||||| | ||| | | || | ||| || |
Db      4126 TGGGCCCAGCTGTGACGTGTCTGTTCACAGGGCACTTCTGGCTTCTTCTGCCCCAGCAC 4067

Qy      711 ATGCCCTTGTCAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTC 770
      |||| ||||| ||||| || | || | || ||| ||| |||| |
Db      4066 CCATCCTTGCCAAAATGGAGGTGTCTTCAAACCCACAGGGCTCCTGCAGCTGCCCCC 4007

Qy      771 TGGCTGGATGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTG 830
      ||||| ||||| || || | || |||| |||| ||| ||||
Db      4006 TGGCTGGATGGGCACCATCTGCTCCCTGCCCTGCCAGAGGGCTTTCACGGACCCAACTG 3947

Qy      831 TTCCAAGAATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCA 890
      |||| |||| | |||| || || || || |||| ||| || || ||
Db      3946 CTCCCAGGAATGTCGCTGCCACAACGGCGGCCTCTGTGACCGATTCACTGGGCAGTGCCG 3887

Qy      891 TTGCAGTCCAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGG 950
      ||| ||| || |||| |||| ||||| || |||| || || ||
Db      3886 CTGCGCTCCGGGTTACACTGGGGATCGGTGCCGGGAGGAGTGCCCGGTGGGCCGCTTTGG 3827

Qy      951 CGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGG 1010
      ||||| ||||| ||| || | | | | | || || || | |||
Db      3826 GCAGGACTGTGCTGAGACGTGCGACTGCGCCCCGGACGCCCGTTGCTTCCCGGCCAACGG 3767

Qy      1011 CGCATGCCTCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGA 1070
      ||||| || || || || |||| |||| || ||||| | |||| || || ||
Db      3766 CGCATGTCTGTGCGAACACGGCTTCACTGGGGACCGCTGCACGGATCGCCTTGCCCCGA 3707

Qy      1071 GGGGCTCTACGGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAG 1130
```


Db	3706	CGGCTTCTACGGTCTCAGCTGCCAGGCCCCCTGCACCTGCGACCGGGAGCACAGCCTCAG	3647
Qy	1131	CTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAA	1190
Db	3646	CTGCCACCCGATGAACGGGGAGTGCTCCTGCCTGCCGGGTGGGCGGGCCTCCACTGCAA	3587
Qy	1191	TGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAA	1250
Db	3586	CGAGAGCTGCCCCGAGGACACGCATGGGCCAGGGTGCCAGGAGCACTGTCTCTGCCTGCA	3527
Qy	1251	TGGGGCAGACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAAT	1310
Db	3526	CGGTGGCGTCTGCCAGGCTACCAGCGGCCTCTGTCACTGCGCGCCGGGTACACGGGCCC	3467
Qy	1311	TGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGG	1370
Db	3466	TCACTGTGCTAGTCTTTGTCTCCTGACACCTACGGTGTCAACTGTTCTGCACGCTGCTC	3407
Qy	1371	CTGTAAAAATGATGCACTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTG	1430
Db	3406	ATGTGAAAATGCCATCGCCTGCTCACCCATCGACGGCGAGTGCCTCTGCAAGGAAGGTTG	3347
Qy	1431	GCACGGGGTGGACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACCT	1490
Db	3346	GCAGCGTGGTAACTGCTCTGTGCCCTGCCCACCCGGAACCTGGGGCTTCAGTTGCAATGC	3287
Qy	1491	AACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGC	1550
Db	3286	CAGCTGCCAGTGTGCCATGAGGCAGTCTGCAGCCCCCAAACCTGGAGCCTGTACCTGCAC	3227
Qy	1551	ACCTGGATGGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAA	1610
Db	3226	CCCTGGGTGGCATGGGGCCCACTGCCAGCTGCCCTGTCCGAAGGGGCAGTTTGGAGAAGG	3167
Qy	1611	CTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTG	1670
Db	3166	TTGTGCCAGTCGCTGTGACTGTGACCACTCTGATGGCTGTGACCCTGTTTCATGGACGCTG	3107
Qy	1671	CCGCTGCCCTCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTG	1730
Db	3106	TCAGTGCCAGGCTGGCTGGATGGGTGCCCGCTGCCACCTGTCCTGCCCTGAGGGCTTATG	3047
Qy	1731	GGGCCCCAACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCTGATGA	1790
Db	3046	GGGAGTCAACTGTAGCAACACCTGCACCTGCAAGAATGGGGGCACCTGTCTCCCTGAGAA	2987
Qy	1791	TGGCATCTGCGAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCC	1850
Db	2986	TGGCAACTGCGTGTGTGCACCCGGATTCCGGGGCCCCCTCCTGCCAGAGATCCTGTCAGCC	2927
Qy	1851	TGGTTTTTATGGGCATCGCTGCAGCCAGACATGCCACAGTGCCTTCACAGCAGCGGGCC	1910
Db	2926	TGGCCGCTATGGCAAACGCTGTG-----TGCCCTGCAAGTGCCTAACCCTCCTTC	2875
Qy	1911	CTGCCACCACATACCGGCCTGTGTGACTGCTTGCCCTGGCTTCACAGGCGCCCTCTGCAA	1970

Db 2874 TGCCACCCCTTCGAACGGGACCTGCTACTGCCTGGCTGGCTGGACAGGCCCCGACTGCTC 2815

Qy 1971 TGAAGTGTGTCCCAAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAA 2030
 | | | | | | | | | | | | | | | | | | | | | |

Db 2814 CCAGCCATGCCCTCCAGGACACTGGGGAGAAAACTGTGCCAGACCTGCCAATGTCACCA 2755

Qy 2031 CAACGGAACCTGTAACCCCATTTGACAGATCTTGTCAAGTGTACCCCGTTGGATTGGCAG 2090
 | | | | | | | | | | | | | | | | | | | | | |

Db 2754 TGGTGGGACCTGCCATCCCCAGGATGGGAGCTGTATCTGCCCCCTAGGCTGGACTGGACA 2695

Qy 2091 TGAAGTGTGTCCCAAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAA 2150
 | | | | | | | | | | | | | | | | | | | | | |

Db 2694 CCACTGCTTAGAAGGCTGCCCTCTGGGGACATTTGGTGTCAACTGCTCCAGCCATGCCA 2635

Qy 2151 CTGCCATAATGGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGG 2207
 | | | | | | | | | | | | | | | | | | | | | |

Db 2634 GTGTGGTCTTGGAGAAAAGTGCCACCCAGAGACTGGGGCCTGTGTATGTCCCCCAGG 2578

RESULT 3

US-09-130-491-10

; Sequence 10, Application US/09130491

; Patent No. 6416974

; GENERAL INFORMATION:

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Goodearl, Andrew D.J.

; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83

; FILE REFERENCE: 09404/041001

; CURRENT APPLICATION NUMBER: US/09/130,491

; CURRENT FILING DATE: 1998-08-07

; EARLIER APPLICATION NUMBER: US 60/058,108

; EARLIER FILING DATE: 1997-09-05

; EARLIER APPLICATION NUMBER: US 60/054,961

; EARLIER FILING DATE: 1997-08-06

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 10

; LENGTH: 1578

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(1578)

; OTHER INFORMATION: n = A,T,C or G

US-09-130-491-10

Query Match 4.8%; Score 165.8; DB 4; Length 1578;

Best Local Similarity 95.5%; Pred. No. 6e-41;

Matches 191; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

Qy 3227 AAGTTGAACCTACAGTGAGTGTGTGT--CCAAGGAGTATTTCAGCAATAATGGGCGTCTCTC 3284
 | | | | | | | | | | | | | | | | | | | | | |

Db 25 AAGTGAACCTAACAGTGAGTGTGTGTTCCTCAAGGAGTATTTCAGCAATAATGGGCGTCTNTC 84

Qy 3285 CC-AGGATCCATATGACCTCCCAAAGAACAGTCACATCCCTTGTCAATTATGACCTGCTGC 3343
 | | | | | | | | | | | | | | | | | | | | | |

Db 85 CCAAGGATCCATATGACCTCCCAAAGAACAGTCACATCCCTTGTCAATTATGACCTGCTGC 144

; TYPE: DNA
; ORGANISM: Mouse
US-09-312-283C-255

Query Match 4.6%; Score 156; DB 4; Length 1464;
Best Local Similarity 51.2%; Pred. No. 6.5e-38;
Matches 419; Conservative 0; Mismatches 390; Indels 9; Gaps 2;

```
Qy      491 CCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGC 550
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      9   CGGGAGCCTGCTACTGCCCTGCTGGGTTTCCTTGGGGCCGACTGTAGCCTTGCCTGTCCAC 68

Qy      551 AGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCG 610
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      69   AGGGTCGCTTCGGCCCCAGCTGTGCCACGTGTGTACATGCGGGCAAGGGGCGGCATGTG 128

Qy      611 ACCACGTCACGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTCTGTGAGGATC 670
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      129 ACCCAGTGTGCGGGGACTTGCATCTGTCTCCCGGGAAGACGGGAGGCCATTGTGAGCGCG 188

Qy      671 TTTGTCCTCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAATGGAG 730
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      189 GCTGTCCCCAGGACCGGTTTGGCAAGGGCTGTGAACACAAGTGTGCCTGCAGGAATGGGG 248

Qy      731 GAGTGTGTATCACGTCACCTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGT 790
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      249 GCCTGTGTATGCTACCAATGGCAGCTGCTCCTGCCCCCTGGGCTGGATGGGGCCACACT 308

Qy      791 GTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCC 850
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      309 GTGAGCACGCCTGCCCTGCTGGGCGCTATGGTGTGCTGCCTGCCTCCTGGAGTGTTCCTGTC 368

Qy      851 ATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAG 910
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      369 AGAACAATGGCAGCTGTGAGCCACCTCCGGCGCTTGCTCTGTGGCCCTGGCTTCTATG 428

Qy      911 GGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCT 970
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      429 GTCAAGCTTGTGAAGACACCTGCCCTGCCGGCTTCCATGGATCTGGTTGCCAGAGAGTTT 488

Qy      971 GCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAG 1030
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      489 GCGAGTGTCAACAGGGCGCTCCCTGTGACCCTGTCAGTGGCCGGTGCCTCTGCCCTGCTG 548

Qy      1031 GCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTTGAGGGGCTCTACGGCATCAAAT 1090
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      549 GCTTCCGTGGCCAGTTCTGCGAGAGGGG---GTGCAAGCCAGGCTTTTTTGGAGATGGCT 605

Qy      1091 GTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTACCCCATGTCTGGAG 1150
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      606 GCCTGCAGCAGTGTAACTGCCCCACGGGTGTGCC-----CTGTGATCCCATCAGCGGCC 659

Qy      1151 AGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGAT 1210
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      660 TCTGCCTTTGCCACCAGGGCGCGCAGGAACCACATGTGACCTAGATTGCAGAAGAGGCC 719

Qy      1211 TCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTG 1270
```

```

      || ||| | || | ||| || ||||| ||||| |||
Db      720 GCTTTGGGCCGGGCTGTGCCCTGCGCTGTGATTGTGGGGGTGGGGCTGACTGCGACCCCA 779

Qy      1271 TGA CTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGA 1308
      || ||| |||| |||| | || |||
Db      780 TCAGTGGGCAGTGCCACTGTGTGGACAGCTACACGGGA 817

```

RESULT 6

US-09-312-283C-73

; Sequence 73, Application US/09312283C

; Patent No. 6573095

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated from Skin Cells

; TITLE OF INVENTION: and Methods for Their Use

; FILE REFERENCE: 11000.1011c2

; CURRENT APPLICATION NUMBER: US/09/312,283C

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 425

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 73

; LENGTH: 1635

; TYPE: DNA

; ORGANISM: Mouse

US-09-312-283C-73

Query Match 4.6%; Score 156; DB 4; Length 1635;

Best Local Similarity 51.2%; Pred. No. 7e-38;

Matches 419; Conservative 0; Mismatches 390; Indels 9; Gaps 2;

```

Qy      491 CCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGC 550
      | ||| || ||| |||| |||| || |||| || || |||| ||||
Db      9   CGGGAGCCTGCTACTGCCCTGCTGGGTTCCCTTGGGGCCGACTGTAGCCTTGCCCTGTCCAC 68

Qy      551 AGGGCACCTATGGTAACGACTGTGCATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCG 610
      |||| || || || |||| || || || || || || || || ||
Db      69 AGGGTCGCTTCGGCCCCAGCTGTGCCACGCTGTGTACATGCGGGCAAGGGGCGGCATGTG 128

Qy      611 ACCACGTCACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTCTGTGAGGATC 670
      ||| || |||| || || || || || || || || || || ||
Db      129 ACCCAGTGTGCGGGGACTTGCATCTGTCTCCCGGGAAGACGGGAGGCCATTGTGAGCGCG 188

Qy      671 TTTGTCCTCCTGGTAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAAATGGAG 730
      ||||| | | ||| |||| || | || | || |||| |
Db      189 GCTGTCCCCAGGACCGGTTTGGCAAGGGCTGTGAACACAAGTGTGCCTGCAGGAATGGGG 248

Qy      731 GAGTGTGTGCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGT 790
      | ||||| || || || |||| |||| ||||| || || |
Db      249 GCCTGTGTGCATGCTACCAATGGCAGCTGCTCCTGCCCCCTGGGCTGGATGGGGCCACACT 308

Qy      791 GTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCC 850

```

```

      |||  ||  | ||||  |  || |||  |||      |||  ||  ||  ||  ||  |
Db      309 GTGAGCACGCCTGCCCTGCTGGGCGCTATGGTGCTGCCTGCCTCCTGGAGTGTTCCTGTC 368

Qy      851 ATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAG 910
      |  ||  ||  | ||||  |  ||  |||  ||  |  ||  ||  ||  ||  |
Db      369 AGAACAATGGCAGCTGTGAGCCACCTCCGGCGCTTGCCCTCTGTGGCCCTGGCTTCTATG 428

Qy      911 GGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCT 970
      |  ||  ||  | ||  ||  || |||  ||  |  ||||  |  ||  |
Db      429 GTCAAGCTTGTGAAGACACCTGCCCTGCCGGCTTCCATGGATCTGGTTGCCAGAGAGTTT 488

Qy      971 GCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAG 1030
      ||  ||||  |  ||  ||  || |||  ||  ||  ||  ||  ||||  ||  |
Db      489 GCGAGTGTCAACAGGGCGCTCCCTGTGACCTGTGAGTGGCCGGTGCCTCTGCCCTGCTG 548

Qy      1031 GCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTACGGCATCAAAT 1090
      ||||  ||||  ||  ||||  |  |||  ||  ||  ||  ||  ||  |
Db      549 GCTTCCGTGGCCAGTTCTGCGAGAGGGG---GTGCAAGCCAGGCTTTTTTGGAGATGGCT 605

Qy      1091 GTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTACCCCATGTCTGGAG 1150
      |  |  ||||  ||||  |  ||  |  ||||  ||||  ||  ||
Db      606 GCCTGCAGCAGTGTAAGTACCCACGGGTGTGCC-----CTGTGATCCCATCAGCGGCC 659

Qy      1151 AGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGAT 1210
      ||  |||  ||  ||  |  ||||  |  |||  |  ||  ||
Db      660 TCTGCCTTTGCCACCAGGGCGCGCAGGAACCACATGTGACCTAGATTGCAGAAGAGGCC 719

Qy      1211 TCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTG 1270
      ||  |||  |  ||  |  ||  ||  ||  ||||  ||||  ||
Db      720 GCTTTGGGCCGGGCTGTGCCCTGCGCTGTGATTGTGGGGGTGGGGCTGACTGCGACCCCA 779

Qy      1271 TGAAGTGCACCTGTGCCCCAGGATTCAAAGGA 1308
      |  |||  ||||  ||||  |  ||  |||  ||
Db      780 TCAGTGGGCAGTGCCACTGTGTGGACAGCTACACGGGA 817

```

RESULT 7

US-09-188-930-73

; Sequence 73, Application US/09188930A

; Patent No. 6150502

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James Greg

; TITLE OF INVENTION: Compositions Isolated From Skin Cells

; TITLE OF INVENTION: and Methods For Their Use

; FILE REFERENCE: 11000.1011c1

; CURRENT APPLICATION NUMBER: US/09/188,930A

; CURRENT FILING DATE: 1998-11-09

; NUMBER OF SEQ ID NOS: 348

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 73

; LENGTH: 1633

; TYPE: DNA

; ORGANISM: mouse

; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1608)...(1608)
US-09-188-930-73

Query Match 4.5%; Score 155.2; DB 3; Length 1633;
Best Local Similarity 51.0%; Pred. No. 1.2e-37;
Matches 417; Conservative 2; Mismatches 390; Indels 9; Gaps 2;

```
Qy      491 CCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGC 550
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      9 CGGGAGCCTGCTACTGCCCTGCTGGGTTCCCTTGGGGCCGACTGTAGCCTTGCCTGTCCAC 68

Qy      551 AGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCG 610
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      69 AGGGTCGCTTCGGCCCCAGCTGTGCCACGCTGTGTACATGCGGGCAAGGGGCGGCATGTG 128

Qy      611 ACCACGTCACGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTCTGTGAGGATC 670
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      129 ACCCAGTGTCGGGGACTTGCATCTGTCTCCCGGAAGACGGGAGGCCATTGTGAGCGCG 188

Qy      671 TTTGTCTCTCGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAAATGGAG 730
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      189 GCTGTCCCAGGACCGGTTTGGCAAGGGCTGTGAACACAAGTGTGCCTGCAGGAATGGGG 248

Qy      731 GAGTGTGTCATCACGTCACCTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGT 790
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      249 GCCTGTGTCATGCTACCAATGGCAGCTGCTCCTGCCCTTGGCTGKATGGGGCCACACT 308

Qy      791 GTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAAATGCCAGTGCC 850
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      309 GTGAGCACGCCTGCCCTGCTGGGCGCTATGGTGCTGCCTGCCTCCTGGAGTGTTCCTGTC 368

Qy      851 ATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGATACACAG 910
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      369 AGAACAATGGCAGCTGTGAGCCACCTCCGGCGCTTGCCTCTGTGGCCCTGGCTTCTATG 428

Qy      911 GGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCT 970
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      429 GTCAAGCTTGTGAAGACACCTGCCCTGCCGGCTTCCATGGATCTGGTTGCCAGAGAGTTT 488

Qy      971 GCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAG 1030
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      489 GCGAGTGTCAACAGGGCGCTCCCTGTGACCCTGTCAGTGGCCGGTGCCTCTGCCCTGCTG 548

Qy      1031 GCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTACGGCATCAAAT 1090
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      549 GCTTCCGTGGCCAGTTCTGCGAGAGGGG---GTGCAAGCCAGGCTTTTTTGGAGATGGCT 605

Qy      1091 GTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTCACCCCATGTCTGGAG 1150
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      606 GCCTGCAGCAGTGTAAGTACCCACGGGTGTGCC-----CTGTGATCCCATCAGCGGCC 659

Qy      1151 AGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGAT 1210
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      660 TCTGCCTTTGCCACCAGGGCGCGCAGGAACACATGTGACCTAGATTGCAGAAGARGCC 719
```


Qy 1211 TCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTG 1270
 || ||| | || | || | || ||||| ||||| |||
 Db 720 GCTTTGGGCCGGGCTGTGCCCTGCGCTGTGATTGTGGGGTGGGGCTGACTGCGACCCCA 779
 Qy 1271 TGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGA 1308
 | | ||| |||| | || | || |||
 Db 780 TCAGTGGGCAGTGCCACTGTGTGGACAGCTACACGGGA 817

RESULT 8

US-09-833-381-1076

; Sequence 1076, Application US/09833381

; Patent No. 6672186

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 1076

; LENGTH: 393

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-833-381-1076

Query Match 3.9%; Score 134.4; DB 4; Length 393;
 Best Local Similarity 61.9%; Pred. No. 1.4e-31;
 Matches 213; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

Qy 781 GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA 840
 || ||||| |||| | || | || | ||||| ||| ||
 Db 28 GGAGCAGTGTGTGCCAGCCCTGCCACCAGGGACATTTGGCCAGAAGTGCAGCCAGGAT 87
 Qy 841 TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA 900
 || | |||| | |||| | |||| | || || || || || || ||
 Db 88 TGTCTTGCCACCATGGAGGGCAGTGTGACCACGTGACTGGACAGTGCCACTGTACAGCT 147
 Qy 901 GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTTGGGACCTATGGCGTTCTCTGT 960
 ||||| |||| | |||| | || || || || || || || || || ||
 Db 148 GGATACATGGGGGACAGGTGCCAAGAGGAGTGCCCTTCGGGTCTTCGGCTTCCAGTGC 207
 Qy 961 GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTTACCACGTGAGCGGCGCATGCCTC 1020
 | || || | || || || || || || || || | || || || ||
 Db 208 TCACAGCGCTGTGACTGCCACAATGGGGGGCAGTGTTACCCACCACGGGTGCCTGCGAG 267
 Qy 1021 TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGCTCTAC 1080
 |||| | || | || | |||| | | || |||| || |||| || |
 Db 268 TGTGAGCCTGGCTACAAGGGCCACGCTGCCAGGAGCGACTGTGCCCGGAGGGCCTGCAT 327
 Qy 1081 GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAC 1124
 ||| || | | || || || || || || || || || || || ||
 Db 328 GGCCCAGGCTGCACCTGCCCTGCCCTGTGACGCTGACAACAC 371

RESULT 9

US-09-621-976-1192

; Sequence 1192, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Jobert, S.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621,976
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 19335
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 1192
 ; LENGTH: 553
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 331..552
 US-09-621-976-1192

Query Match 1.9%; Score 66.2; DB 4; Length 553;
 Best Local Similarity 60.1%; Pred. No. 3.6e-10;
 Matches 110; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 2996 ACAGAAGCTATATGGGAAAATCCTTAAAGACCTGGGAAAGAATTCTGAATATAATTCAA 3055
 ||| | | |||| | | | |||| | | | ||| |
 Db 347 ACACATACATTATGGACAAAGGCTTCAAAGATTACATGAAAGAATCCGTGTGCAGTTCTA 406
 Qy 3056 GTAAGTCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTAC 3115
 ||| || ||| | | || |||| || |||| |||| |||| ||
 Db 407 GTACTTGTTTCCTTGAATAGCAGTGAAAACCCCTTACGCCACAATTAAGGACCCACCCATCC 466
 Qy 3116 TTATCCCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCAGCAAGAGATT 3175
 ||| || |||| | | |||| || |||| |||| | | | |
 Db 467 TCACCTGCAAGCTTCCAGAAAAGCAGCTATGTAGAAATGAAGTCGCCTGTGCACATGGGGT 526
 Qy 3176 CCC 3178
 ||
 Db 527 CTC 529

RESULT 10

US-08-323-474-1

; Sequence 1, Application US/08323474
 ; Patent No. 5447860
 ; GENERAL INFORMATION:
 ; APPLICANT: Ziegler, Steven F.
 ; TITLE OF INVENTION: NOVEL TYROSINE KINASE
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Immunex Corporation
 ; STREET: 51 University Street
 ; CITY: Seattle

```

; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323,474
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/905,600
; FILING DATE: 26-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2609
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 149..3523
US-08-323-474-1

```

```

Query Match          1.6%; Score 55.2; DB 1; Length 4138;
Best Local Similarity 60.8%; Pred. No. 4e-06;
Matches 90; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```

```

Qy      1131 CTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAA 1190
        ||||| |||   |   | |||| ||||   | |||||   || ||   || ||
Db      985 CTGTCTCCCTGACCCCTATGGGTGTTCTGTGCCACAGGCTGGAAGGGTCTGCAGTGCAA 1044

Qy      1191 TGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAA 1250
        |||  ||||   ||||| || |||||   || |||   |||   | ||||| ||| ||
Db      1045 TGAAGCATGCCACCCTGGTTTTTACGGGCCAGATTGTAAGCTTAGGTGCAGCTGCAACAA 1104

Qy      1251 TGGGGCAGACTGTGACAGTGTGACTGGA 1278
        |||||   ||||| | |   |||
Db      1105 TGGGGAGATGTGTGATCGCTTCCAAGGA 1132

```

RESULT 11

PCT-US93-06093-1

; Sequence 1, Application PC/TUS9306093

; GENERAL INFORMATION:

; APPLICANT: Ziegler, Steven F.

```

; TITLE OF INVENTION: NOVEL TYROSINE KINASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06093
; FILING DATE: 19930625
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/905,600
; FILING DATE: 26-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2609
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4138 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 149..3523
PCT-US93-06093-1

```

```

Query Match          1.6%; Score 55.2; DB 5; Length 4138;
Best Local Similarity 60.8%; Pred. No. 4e-06;
Matches 90; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```

```

Qy      1131 CTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAA 1190
          ||||| |||      |      | |||| ||||      | |||||      || || | || ||
Db      985 CTGTCTCCCTGACCCCTATGGGTGTTCTGTGCCACAGGCTGGAAGGGTCTGCAGTGCAA 1044

Qy      1191 TGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAA 1250
          |||  ||||  ||||| || |||||  || |||  |||  | ||||| ||| ||
Db      1045 TGAAGCATGCCACCCTGGTTTTTACGGGCCAGATTGTAAGCTTAGGTGCAGCTGCAACAA 1104

Qy      1251 TGGGGCAGACTGTGACAGTGTGACTGGA 1278
          |||||  |||||  |  |  |||
Db      1105 TGGGGAGATGTGTGATCGCTTCCAAGGA 1132

```

RESULT 12

US-08-220-240A-4

; Sequence 4, Application US/08220240A

; Patent No. 5955291

; GENERAL INFORMATION:

; APPLICANT: Alitalo, Kari

; APPLICANT: Matikainen, Marja-Terttu

; APPLICANT: Partanen, Juha

; APPLICANT: Makela, Tomi

; APPLICANT: Korhonen, Jaana

; TITLE OF INVENTION: ANTIBODIES RECOGNIZING TIE RECEPTOR

; TITLE OF INVENTION: TYROSINE KINASE AND USES THEREOF

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 233 South Wacker Drive/6300 Sears Tower

; CITY: Chicago

; STATE: Illinois

; COUNTRY: Unites States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/220,240A

; FILING DATE: 29-MAR-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/FI93/00006

; FILING DATE: 08-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/817,800

; FILING DATE: 09-JAN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/167,453

; FILING DATE: 15-DEC-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Gass, David A.

; REGISTRATION NUMBER: 38,153

; REFERENCE/DOCKET NUMBER: 29151/31958

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 474-6300

; TELEFAX: (312) 474-0448

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3845 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 37..3450

US-08-220-240A-4

Query Match 1.6%; Score 54; DB 2; Length 3845;
 Best Local Similarity 51.4%; Pred. No. 9e-06;
 Matches 179; Conservative 0; Mismatches 160; Indels 9; Gaps 2;

```

Qy      1709 GCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC---TACTGTAAAA 1765
          | | |||| | || ||||| || || || | || |
Db      674 GGGGTTGTGGGGCTGGGCGCTGGGGGCCAGGCTGTACCAAGGAGTGCCAGGTTGCCTAC 733

Qy      1766 ATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGCTTCCGAGGCA 1825
          |||| | | || | |||| || | | | || |||| ||||
Db      734 ATGGAGGTGTCTGCCACGACCATGACGGCGAATGTGTATGCCCCCTGGCTTCACTGGCA 793

Qy      1826 CCACTTGTGAGAGGATCTGCTCCCCTGGTTTATGGGCATCGCTGCAGCCAGACATGCC 1885
          || ||| | | |||| || || |||| |||| || ||||
Db      794 CCCGCTGTGAACAGGCCTGCAGAGAGGGCCGTTTGGGCAGAGCTGCCAGGAGCAGTGCC 853

Qy      1886 CACAGTGCGTTACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGTGACTGCTTGC 1945
          || | || | | | | | | | | | | | | | |
Db      854 CAGGCATATCAGGCTGCCGGGGCCCTCACCTTCTGCCTCCCAGACCCCTATGGCTGCTCTT 913

Qy      1946 -----CTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTGGGA 1999
          ||||| | ||| || ||| | |||| || || || |||||
Db      914 GTGGATCTGGCTGGAGAGGAAGCCAGTGCCAAGAAGCTTGTGCCCCTGGTCATTTGGGG 973

Qy      2000 AAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACC 2047
          | || | || || || || || || || ||
Db      974 CTGATTGCCGACTCCAGTGCCAGTGTGAGAATGGTGGCACTTGTGACC 1021
  
```

RESULT 13

US-07-934-393B-1

; Sequence 1, Application US/07934393B

; Patent No. 5466596

; GENERAL INFORMATION:

; APPLICANT: BREITMAN, MARTIN L.

; APPLICANT: DUMONT, DANIEL

; APPLICANT: GRADWOHL, GERARD G.

; TITLE OF INVENTION: TISSUE SPECIFIC TRANSCRIPTIONAL

; TITLE OF INVENTION: REGULATORY ELEMENT

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BERESKIN & PARR

; STREET: 40 King Street West

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5H 3Y2

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/934,393B

; FILING DATE: 25-AUG-1992

; CLASSIFICATION: 435

```

; ATTORNEY/AGENT INFORMATION:
;   NAME:  Kurdydyk, Linda M.
;   REGISTRATION NUMBER:  34,971
;   REFERENCE/DOCKET NUMBER:  3153-64
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (416) 354-7311
;   TELEFAX:  (416) 361-1398
; INFORMATION FOR SEQ ID NO:  1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  4175 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   HYPOTHETICAL:  NO
;   ANTI-SENSE:  NO
;   FRAGMENT TYPE:  N-terminal
;   ORIGINAL SOURCE:
;     ORGANISM:  Mus pahari
;     STRAIN:  CD-1
;     DEVELOPMENTAL STAGE:  Embryo
;     TISSUE TYPE:  Heart
;   IMMEDIATE SOURCE:
;     CLONE:  tek
;   POSITION IN GENOME:
;     CHROMOSOME/SEGMENT:  4
;     MAP POSITION:  Between the brown and pmv-23 loci
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  124..3477
;     OTHER INFORMATION:  /function= "putative transmembrane
;     OTHER INFORMATION:  receptor"
;     OTHER INFORMATION:  /product= "tyrosine kinase"
;     OTHER INFORMATION:  /gene= "tek"
;     OTHER INFORMATION:  /standard_name= "tyrosine kinase receptor protein"
US-07-934-393B-1

```

```

Query Match          1.6%;  Score 54;  DB 1;  Length 4175;
Best Local Similarity  50.6%;  Pred. No. 9.6e-06;
Matches 166;  Conservative  0;  Mismatches 150;  Indels  12;  Gaps  1;

```

```

Qy      716 CTTGTCAAAATGGAGGAGTGTGTTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCT 775
        |||| | || |||| | || | | | | | | | | | | | |
Db      800 CTTGCAAGAACAATGGAGTCTGCCATGAAGATACCGGGAATGCATTTGCCCTCCTGGGT 859

Qy      776 GGATGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCC 835
        |||| | | ||| | ||| | | |||| | | |||
Db      860 TTATGGGGAGAACATGTGAGAAAGCTTGTGAGCCGCACACATTTGGCAGGACCTGTAAAG 919

Qy      836 AAGAATGCCAGTGCCATAATGGAGGGACGTGTGATG-----CTGCCACAGGCC 883
        || || | | | ||| | | || | || | | | |
Db      920 AAAGGTGTAGTGGACCAAGAAGGATGCAAGTCTTATGTGTTCTGTCTCCAGACCCTTACG 979

Qy      884 AATGTCATTGCAGTCCAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGA 943
        ||| || |||| | | ||| | |||| | | | | |||
Db      980 GGTGTTCTGTGCCACAGGCTGGAGGGGGTTGCAGTGCAATGAAGCATGCCCATCTGGTT 1039

```

Qy 944 CCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACG 1003
 ||| || |||| | |||| ||| ||| | ||| ||| |
 Db 1040 ACTACGGACCAGACTGTAAGCTCAGGTGCCACTGTACCAATGAAGAGATATGTGATCGGT 1099
 Qy 1004 TGAGCGGCGCATGCCTCTGTGAAGCAGG 1031
 | || | ||| | |
 Db 1100 TCCAAGGATGCCTCTGCTCTCAAGGATG 1127

RESULT 14

US-08-278-089A-1

; Sequence 1, Application US/08278089A

; Patent No. 5681714

; GENERAL INFORMATION:

; APPLICANT: Breitman, Martin L.

; APPLICANT: Rossant, Janet

; APPLICANT: Dumont, Daniel J.

; APPLICANT: Yamaguchi, Terry P.

; TITLE OF INVENTION: No. 5681714e1 Receptor Tyrosine Kinase

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bereskin & Parr

; STREET: 40 King Street West

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5H 3Y2

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/278,089A

; FILING DATE: 20-JUL-1994

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Kurdydyk, Linda M.

; REGISTRATION NUMBER: 34,971

; REFERENCE/DOCKET NUMBER: 3153-111

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 364-7311

; TELEFAX: (416) 361-1398

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4175 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

; ORGANISM: Mus musculus

; STRAIN: CD-1

; DEVELOPMENTAL STAGE: Embryo


```

; TISSUE TYPE: Heart
; IMMEDIATE SOURCE:
; CLONE: Tek
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 4
; MAP POSITION: Between the brown and pmv-23 loci
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 124..3478
US-08-278-089A-1

```

```

Query Match          1.6%; Score 54; DB 1; Length 4175;
Best Local Similarity 50.6%; Pred. No. 9.6e-06;
Matches 166; Conservative 0; Mismatches 150; Indels 12; Gaps 1;

```

```

Qy      716 CTTGTCAAATGGAGGAGTGTGTCAACGTCCTGGAGAATGCTCTTGCCCTTCTGGCT 775
        |||| | || |||| | || | | || || |||| | |||| | || |
Db      800 CTTGCAAGAACAATGGAGTCTGCCATGAAGATAACGGGGAATGCATTTGCCCTCCTGGGT 859

Qy      776 GGATGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCC 835
        |||| | | || | | || | | || | | || | | || |
Db      860 TTATGGGGAGAACATGTGAGAAAGCTTGTGAGCCGCACACATTTGGCAGGACCTGTAAAG 919

Qy      836 AAGAATGCCAGTGCCATAATGGAGGGACGTGTGATG-----CTGCCACAGGCC 883
        || || | | | || | || | || | || | || | |
Db      920 AAAGGTGTAGTGGACCAGAAGGATGCAAGTCTTATGTGTTCTGTCTCCAGACCCTTACG 979

Qy      884 AATGTCATTGCAGTCCAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGA 943
        ||| || |||| | | || | |||| | || | || | |||
Db      980 GGTGTTCTGTGCCACAGGCTGGAGGGGCTTGCAGTGCAATGAAGCATGCCCATCTGGTT 1039

Qy      944 CCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACG 1003
        ||| || |||| | |||| || | || | || | || | |
Db      1040 ACTACGGACCAGACTGTAAGCTCAGGTGCCACTGTACCAATGAAGAGATATGTGATCGGT 1099

Qy      1004 TGAGCGGCGCATGCCTCTGTGAAGCAGG 1031
        | || | || | || |
Db      1100 TCCAAGGATGCCTCTGCTCTCAAGGATG 1127

```

RESULT 15

US-08-838-957A-1

; Sequence 1, Application US/08838957A

; Patent No. 5998187

; GENERAL INFORMATION:

; APPLICANT: Breitman, Martin L.

; APPLICANT: Rossant, Janet

; APPLICANT: Dumont, Daniel J.

; APPLICANT: Yamaguchi, Terry P.

; TITLE OF INVENTION: No. 5998187e1 Receptor Tyrosine Kinase

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bereskin & Parr

; STREET: 40 King Street West

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

```

; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,957A
; FILING DATE: 23-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurdydyk, Linda M.
; REGISTRATION NUMBER: 34,971
; REFERENCE/DOCKET NUMBER: 3153-212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; STRAIN: CD-1
; DEVELOPMENTAL STAGE: Embryo
; TISSUE TYPE: Heart
; IMMEDIATE SOURCE:
; CLONE: Tek
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 4
; MAP POSITION: Between the brown and pmv-23 loci
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 124..3478
US-08-838-957A-1

```

```

Query Match          1.6%; Score 54; DB 2; Length 4175;
Best Local Similarity 50.6%; Pred. No. 9.6e-06;
Matches 166; Conservative 0; Mismatches 150; Indels 12; Gaps 1;

```

```

Qy      716 CTTGTCAAATGGAGGAGTGTGTCATCACGTCACCTGGAGAATGCTCTTGCCCTTCTGGCT 775
        |||| | || |||| | | || | | || || |||| | |||| | || |
Db      800 CTTGCAAGAACAATGGAGTCTGCCATGAAGATACCGGGAATGCATTTGCCCTCCTGGGT 859

Qy      776 GGATGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCC 835
        |||| | | |||| | |||| | |||| | || ||||
Db      860 TTATGGGGAGAACATGTGAGAAAGCTTGTGAGCCGCACACATTTGGCAGGACCTGTAAAG 919

Qy      836 AAGAATGCCAGTGCCATAATGGAGGGACGTGTGATG-----CTGCCACAGGCC 883
        || || | | | || | | || | || | || | || |
Db      920 AAAGGTGTAGTGGACCAGAAGGATGCAAGTCTTATGTGTTCTGTCTCCAGACCCTTACG 979

```

```

Qy      884 AATGTCATTGCAGTCCAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGA 943
      |||  ||      |||| | | |||  | |||| | || | || ||  |||
Db      980 GGTGTTCTGTGCCACAGGCTGGAGGGGGTTGCAGTGCAATGAAGCATGCCCATCTGGTT 1039

Qy      944 CCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTTACCACG 1003
      ||| ||      ||||  | ||||| ||| ||| | || | ||| | |
Db      1040 ACTACGGACCAGACTGTAAGCTCAGGTGCCACTGTACCAATGAAGAGATATGTGATCGGT 1099

Qy      1004 TGAGCGGCGCATGCCTCTGTGAAGCAGG 1031
      |  ||      | || | ||| | |
Db      1100 TCCAAGGATGCCTCTGCTCTCAAGGATG 1127

```

Search completed: March 30, 2004, 08:37:02
Job time : 269 secs

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 02:15:41 ; Search time 1109 Seconds
(without alignments)
11491.309 Million cell updates/sec

Title: US-10-092-390-1
Perfect score: 3423
Sequence: 1 atgggtattttctttgaactc.....gcagcagcagcagtgaatga 3423

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%	Query						
No.	Score	Match	Length	DB	ID				Description

	1	3423	100.0	3423	14	US-10-092-390-1	Sequence 1, Appli
	2	2192.8	64.1	3552	14	US-10-365-227-19	Sequence 19, Appl
	3	1760	51.4	1761	14	US-10-092-390-3	Sequence 3, Appli
	4	1425.8	41.7	1448	13	US-10-105-929-9	Sequence 9, Appli
	5	1425.8	41.7	1448	14	US-10-365-227-9	Sequence 9, Appli
	6	640.6	18.7	3114	15	US-10-052-648A-9	Sequence 9, Appli
	7	635.8	18.6	5000	15	US-10-052-648A-7	Sequence 7, Appli
	8	608.4	17.8	3567	10	US-09-796-753-99	Sequence 99, Appl
	9	608.4	17.8	3567	10	US-09-796-753-123	Sequence 123, App
	10	546.2	16.0	2919	15	US-10-052-648A-3	Sequence 3, Appli
	11	546.2	16.0	2919	15	US-10-052-648A-5	Sequence 5, Appli
c	12	540.6	15.8	1970	14	US-10-106-698-1976	Sequence 1976, Ap
	13	528.6	15.4	5036	10	US-09-796-753-113	Sequence 113, App
c	14	524.6	15.3	5197	9	US-09-833-381-1910	Sequence 1910, Ap
	15	511	14.9	3063	15	US-10-052-648A-1	Sequence 1, Appli
	16	427.4	12.5	2220	12	US-10-211-462-68	Sequence 68, Appl
	17	424.8	12.4	1205	14	US-10-084-994-2	Sequence 2, Appli
	18	424.8	12.4	1205	14	US-10-193-109-2	Sequence 2, Appli
	19	424.8	12.4	1205	15	US-10-193-409-2	Sequence 2, Appli
	20	400.8	11.7	5523	15	US-10-191-803-162	Sequence 162, App
	21	395.4	11.6	2700	14	US-10-084-994-5	Sequence 5, Appli
	22	395.4	11.6	2700	14	US-10-193-109-5	Sequence 5, Appli
	23	395.4	11.6	2700	15	US-10-193-409-5	Sequence 5, Appli
	24	387	11.3	5178	12	US-10-312-352-57	Sequence 57, Appl
	25	175	5.1	1686	14	US-10-084-994-6	Sequence 6, Appli
	26	175	5.1	1686	14	US-10-193-109-6	Sequence 6, Appli
	27	175	5.1	1686	15	US-10-193-409-6	Sequence 6, Appli
	28	165.8	4.8	1578	13	US-10-105-929-10	Sequence 10, Appl
	29	165.8	4.8	1578	14	US-10-365-227-10	Sequence 10, Appl
	30	156	4.6	1464	10	US-09-866-050A-255	Sequence 255, App
	31	156	4.6	1464	14	US-10-152-661-255	Sequence 255, App
	32	156	4.6	1627	10	US-09-866-050A-438	Sequence 438, App
	33	156	4.6	1627	14	US-10-152-661-438	Sequence 438, App
	34	154	4.5	1633	10	US-09-866-050A-73	Sequence 73, Appl
	35	154	4.5	1633	14	US-10-152-661-73	Sequence 73, Appl
	36	154	4.5	2569	10	US-09-796-753-115	Sequence 115, App
	37	134.4	3.9	393	9	US-09-833-381-1076	Sequence 1076, Ap
	38	129.6	3.8	1939	9	US-09-804-156-4	Sequence 4, Appli
	39	129.6	3.8	1939	9	US-09-946-633-2	Sequence 2, Appli
	40	129.6	3.8	1939	13	US-10-125-459-2	Sequence 2, Appli
	41	129.6	3.8	1939	13	US-10-067-761-4	Sequence 4, Appli
	42	129.6	3.8	1939	14	US-10-319-519-4	Sequence 4, Appli
	43	89.2	2.6	3408	9	US-09-870-759-133	Sequence 133, App
	44	89.2	2.6	3408	10	US-09-751-708A-133	Sequence 133, App
	45	89.2	2.6	3408	14	US-10-171-581-223	Sequence 223, App

ALIGNMENTS

RESULT 1

US-10-092-390-1

; Sequence 1, Application US/10092390

; Publication No. US20030013865A1

; GENERAL INFORMATION:

; APPLICANT: Yu, Xuanchuan

```

; APPLICANT: Miranda, Maricar
; TITLE OF INVENTION: No. US20030013865A1el Human EGF-Family Proteins and
Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0317-USA
; CURRENT APPLICATION NUMBER: US/10/092,390
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/275,013
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3423
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-092-390-1

```

```

Query Match          100.0%; Score 3423; DB 14; Length 3423;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT 60

Qy     61 GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC 120

Qy    121 TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC 180

Qy    181 ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT 240

Qy    241 CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA 300

Qy    301 AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT 360

Qy    361 CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT 420

Qy    421 GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC 480

Qy    481 AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC 540

Qy    541 CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA 600

```

Db	541		CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	600
Qy	601		GCCACCTGCGACCACGTCACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	660
Db	601		GCCACCTGCGACCACGTCACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	660
Qy	661		TGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Db	661		TGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Qy	721		CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Db	721		CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Qy	781		GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Db	781		GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Qy	841		TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Db	841		TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Qy	901		GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960
Db	901		GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960
Qy	961		GCTGAGACCTGCCAGTGTGTCAACGGAGGGGAAGTGTTACCACGTGAGCGGCGCATGCCTC	1020
Db	961		GCTGAGACCTGCCAGTGTGTCAACGGAGGGGAAGTGTTACCACGTGAGCGGCGCATGCCTC	1020
Qy	1021		TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGGCTCTAC	1080
Db	1021		TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGGCTCTAC	1080
Qy	1081		GGCATCAAATGTGACAAACGGTGTCCTGCCACTTGGAACAACTCATAGCTGTCAACCC	1140
Db	1081		GGCATCAAATGTGACAAACGGTGTCCTGCCACTTGGAACAACTCATAGCTGTCAACCC	1140
Qy	1141		ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Db	1141		ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Qy	1201		TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Db	1201		TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Qy	1261		TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Db	1261		TGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Qy	1321		ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1380
Db	1321		ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAAT	1380
Qy	1381		GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTG	1440

Db	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTG	1440
Qy	1441	GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1500
Db	1441	GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAG	1500
Qy	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1560
Db	1501	TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG	1560
Qy	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1620
Db	1561	CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG	1620
Qy	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1680
Db	1621	CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC	1680
Qy	1681	CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1740
Db	1681	CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC	1740
Qy	1741	TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGC	1800
Db	1741	TGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGC	1800
Qy	1801	GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGACAGGATCTGCTCCCCTGGTTTTTAT	1860
Db	1801	GAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGACAGGATCTGCTCCCCTGGTTTTTAT	1860
Qy	1861	GGGCATCGCTGCAGCCAGACATGCCACAGTGCGTTCACAGCAGCGGGCCCTGCCACCAC	1920
Db	1861	GGGCATCGCTGCAGCCAGACATGCCACAGTGCGTTCACAGCAGCGGGCCCTGCCACCAC	1920
Qy	1921	ATCACCGGCCTGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT	1980
Db	1921	ATCACCGGCCTGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGT	1980
Qy	1981	CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC	2040
Db	1981	CCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACC	2040
Qy	2041	TGTAACCCCATTTGACAGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGACTGCTCT	2100
Db	2041	TGTAACCCCATTTGACAGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGACTGCTCT	2100
Qy	2101	CAACCATGTCCACCTGCCCACTGGGGCCCAAAGTGCATCCACACGTGCAACTGCCATAAT	2160
Db	2101	CAACCATGTCCACCTGCCCACTGGGGCCCAAAGTGCATCCACACGTGCAACTGCCATAAT	2160
Qy	2161	GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC	2220
Db	2161	GGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTC	2220
Qy	2221	TACTGCACTCAGAGATGTCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAA	2280
Db	2221	TACTGCACTCAGAGATGTCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAA	2280

Qy	2281	TGTCAAACGGAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTC	2340
Db	2281	TGTCAAACGGAGCTGACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTC	2340
Qy	2341	ATGGGACGGCACTGTGAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTCGCCAG	2400
Db	2341	ATGGGACGGCACTGTGAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTCGCCAG	2400
Qy	2401	ATATGTGATTGTCTGAACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC	2460
Db	2401	ATATGTGATTGTCTGAACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGC	2460
Qy	2461	CCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAAC	2520
Db	2461	CCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAAC	2520
Qy	2521	AGCTTAAGCCGAACCACTGCTCTCCCTGCTGATTCTACCAGATCGGGGCCATTGCA	2580
Db	2521	AGCTTAAGCCGAACCACTGCTCTCCCTGCTGATTCTACCAGATCGGGGCCATTGCA	2580
Qy	2581	GGCATCATCATTCTTGCTCTAGTTGTTCTCTTCTACTGGCATTGTTTATTATATAGA	2640
Db	2581	GGCATCATCATTCTTGCTCTAGTTGTTCTCTTCTACTGGCATTGTTTATTATATAGA	2640
Qy	2641	CACAAGCAGAAGGGAAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2700
Db	2641	CACAAGCAGAAGGGAAAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGG	2700
Qy	2701	GTCGTCAATGCAGATTATACCATTTCAGGAACCCCTCCTCACAGCAATGGTGGAAACGCT	2760
Db	2701	GTCGTCAATGCAGATTATACCATTTCAGGAACCCCTCCTCACAGCAATGGTGGAAACGCT	2760
Qy	2761	AATAGCCACTACTTCACCAATCCCAGTTACCACACGCTCACCCAGTGTCACACATCCCCT	2820
Db	2761	AATAGCCACTACTTCACCAATCCCAGTTACCACACGCTCACCCAGTGTCACACATCCCCT	2820
Qy	2821	CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	2880
Db	2821	CACGTCAACAACAGGGACAGGATGACTGTCACGAAGTCAAAAAACAATCAACTGTTTGTG	2880
Qy	2881	AATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	2940
Db	2881	AATCTTAAAAATGTGAACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTG	2940
Qy	2941	CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3000
Db	2941	CCGGCTGACTGGAAACATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGA	3000
Qy	3001	AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAC	3060
Db	3001	AGCTATATGGGAAAATCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAAC	3060
Qy	3061	TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3120
Db	3061	TGCTCCCTAAGCAGTTCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATC	3120

Qy 3121 CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCA 3180
 |||
 Db 3121 CCGAAAAGCTCAGAGTGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCCA 3180

Qy 3181 TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA 3240
 |||
 Db 3181 TATGCAGAGATCAATAACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACA 3240

Qy 3241 GTGAGTGTGTCCAAGGAGTATTTCAGCAATAATGGGCGTCTCTCCCAGGATCCATATGAC 3300
 |||
 Db 3241 GTGAGTGTGTCCAAGGAGTATTTCAGCAATAATGGGCGTCTCTCCCAGGATCCATATGAC 3300

Qy 3301 CTCCCAAAGAACAGTCACATCCCTTGTCAATTATGACCTGCTGCCAGTCCGAGACAGTTCA 3360
 |||
 Db 3301 CTCCCAAAGAACAGTCACATCCCTTGTCAATTATGACCTGCTGCCAGTCCGAGACAGTTCA 3360

Qy 3361 TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGCAGTGAA 3420
 |||
 Db 3361 TCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGCAGTGAA 3420

Qy 3421 TGA 3423
 |||
 Db 3421 TGA 3423

RESULT 2

US-10-365-227-19

; Sequence 19, Application US/10365227

; Publication No. US20030143632A1

; GENERAL INFORMATION:

; APPLICANT: McCarthy, Sean A.

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Goodearl, Andrew D.J.

; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING

; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND

OTHER

; TITLE OF INVENTION: USES

; FILE REFERENCE: 07334-323001

; CURRENT APPLICATION NUMBER: US/10/365,227

; CURRENT FILING DATE: 2003-02-12

; PRIOR APPLICATION NUMBER: US/09/802,582

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: US 09/128,709

; PRIOR FILING DATE: 1998-08-04

; PRIOR APPLICATION NUMBER: US 60/054,645

; PRIOR FILING DATE: 1997-08-04

; PRIOR APPLICATION NUMBER: US 09/130,491

; PRIOR FILING DATE: 1998-08-06

; PRIOR APPLICATION NUMBER: US 60/054,966

; PRIOR FILING DATE: 1997-08-06

; PRIOR APPLICATION NUMBER: US 60/058,108

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: US 09/388,280

; PRIOR FILING DATE: 1999-09-01

; PRIOR APPLICATION NUMBER: US 09/388,279

; PRIOR FILING DATE: 1999-09-01

; NUMBER OF SEQ ID NOS: 20

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3552
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1803)
US-10-365-227-19
```

```
Query Match      64.1%; Score 2192.8; DB 14; Length 3552;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2205; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
```

```
Qy      1216 GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT 1275
          |||
Db        1 GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT 60

Qy      1276 GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG 1335
          |||
Db        61 GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG 120

Qy      1336 GGAACCTATGGGATAAACTGTTTCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT 1395
          |||
Db       121 GGAACCTATGGGATAAACTGTTTCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT 180

Qy      1396 CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGA 1455
          |||
Db       181 CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGA 240

Qy      1456 TGTCCCAAGTGGCACATGGGGCTTTGGCTGTAACTTAACATGCCAGTGCCTCAACGGGGGA 1515
          |||
Db       241 TGTCCCAAGTGGCACATGGGGCTTTGGCTGTAACTTAACATGCCAGTGCCTCAACGGGGGA 300

Qy      1516 GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC 1575
          |||
Db       301 GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC 360

Qy      1576 GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC 1635
          |||
Db       361 GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC 420

Qy      1636 CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT 1695
          |||
Db       421 CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT 480

Qy      1696 GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC 1755
          |||
Db       481 GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC 540

Qy      1756 TACTGTAAAAATGGGGCTTCATGCTCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC 1815
          |||
Db       541 TACTGTAAAAATGGGGCTTCATGCTCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC 600

Qy      1816 TTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCTGGTTTTATGGGCATCGCTGCAGC 1875
          |||
Db       601 TTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCTGGTTTTATGGGCATCGCTGCAGC 660
```

Qy	1876	CAGACATGCCCACAGTGC GTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGT	1935
Db	661	CAGACATGCCCACAGTGC GTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGT	720
Qy	1936	GACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT	1995
Db	721	GACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT	780
Qy	1996	GGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTGAC	2055
Db	781	GGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTGAC	840
Qy	2056	AGATCTTGTCA GTGTTACCCCGGTTGGATTGGCAGTGA CTGCTCTCAACCATGTCCACCT	2115
Db	841	AGATCTTGTCA GTGTTACCCCGGTTGGATTGGCAGTGA CTGCTCTCAACCATGTCCACCT	900
Qy	2116	GCCCACTGGGGCCCAA CTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC	2175
Db	901	GCCCACTGGGGCCCAA CTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC	960
Qy	2176	GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA	2235
Db	961	GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA	1020
Qy	2236	TGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACGGAGCT	2295
Db	1021	TGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACGGAGCT	1080
Qy	2296	GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCAC TGGATTTCATGGGACGGCACTGT	2355
Db	1081	GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCAC TGGATTTCATGGGACGGCACTGT	1140
Qy	2356	GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGT CGCCAGATATGTGATTGTCTG	2415
Db	1141	GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGT CGCCAGATATGTGATTGTCTG	1200
Qy	2416	AACAAC TCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA	2475
Db	1201	AACAAC TCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA	1260
Qy	2476	GCGAGATGTGATCAAGCTGGTGT TATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC	2535
Db	1261	GCGAGATGTGATCAAGCTGGTGT TATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC	1320
Qy	2536	AGTACTGCTCTCCCTGCTGATT CCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTT	2595
Db	1321	AGTACTGCTCTCCCTGCTGATT CCTACCACATCGGGGCCATTGCAGGCATCATCATTCTT	1380
Qy	2596	GTCCTAGTTGTTCTCTCTCTACTGGCATTGTTTATTATTTATAGACACAAGCAGAAGGGA	2655
Db	1381	GTCCTAGTTGTTCTCTCTCTACTGGCATTGTTTATTATTTATAGACACAAGCAGAAGGGA	1440
Qy	2656	AAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGGGTCGTCAATGCAGAT	2715
Db	1441	AAGGAATCAAGCATGCCAGCAGTTACCTACACCCCTGCTATGAGGGTCGTCAATGCAGAT	1500

Qy 2716 TATACCATTTCAGGAACCCCTTCCTCACAGCAATGGTGGAAACGCTAATAGCCACTACTTC 2775
 |||
 Db 1501 TATACCATTTCAGGAACCCCTTCCTCACAGCAATGGTGGAAACGCTAATAGCCACTACTTC 1560

Qy 2776 ACCAATCCCAGTTACCACACGCTCACCCAGTGTGCCACATCCCCTCACGTCAACAACAGG 2835
 |||
 Db 1561 ACCAATCCCAGTTACCACACGCTCACCCAGTGTGCCACATCCCCTCACGTCAACAACAGG 1620

Qy 2836 GACAGGATGACTGTACGAAGTCAAAAAACAATCAACTGTTTGTGAATCTTAAAAATGTG 2895
 |||
 Db 1621 GACAGGATGACTGTACGAAGTCAAAAAACAATCAACTGTTTGTGAATCTTAAAAATGTG 1680

Qy 2896 AACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCACTGGGACATTGCCGGCTGACTGGAAA 2955
 |||
 Db 1681 AACCCTGGGAAGAGAGGCCCTGTGGGGGACTGCA-TGGGACATTGCCGGCTGACTGGAAA 1739

Qy 2956 CATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGAAGCTATATGGGAAAA 3015
 |||
 Db 1740 CATGGCGGCTACCTCAACGAGCTCGGTGCTTTTGGACTTGACAGAAGCTATATGGGAAAA 1799

Qy 3016 TCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAACTGCTCCCTAAGCAGT 3075
 |||
 Db 1800 TCCTTAAAAGACCTGGGAAAGAATTCTGAATATAATTCAAGTAACTGCTCCCTAAGCAGT 1859

Qy 3076 TCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATCCCGAAAAGCTCAGAG 3135
 |||
 Db 1860 TCTGAGAACCCATATGCCACTATTAAAGACCCACCTGTACTTATCCCGAAAAGCTCAGAG 1919

Qy 3136 TGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCATATGCAGAGATCAAT 3195
 |||
 Db 1920 TGTGGTTATGTGGAGATGAAATCGCCGGCACGAAGAGATTCCCATATGCAGAGATCAAT 1979

Qy 3196 AACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACAGTGAGTGTGTCCAA 3255
 |||
 Db 1980 AACTCAACTTCAGCCAACAGGAATGTCTATGAAGTTGAACCTACAGTGAGTGTGTCCAA 2039

Qy 3256 GGAGTATTTCAGCAATAATGGGCGTCTCTCCAGGATCCATATGACCTCCCAAAGAACAGT 3315
 |||
 Db 2040 GGAGTATTTCAGCAATAATGGGCGTCTCTCCAGGATCCATATGACCTCCCAAAGAACAGT 2099

Qy 3316 CACATCCCTTGTCAATTATGACCTGCTGCCAGTCCGAGACAGTTCATCCTCCCCTAAGCAA 3375
 |||
 Db 2100 CACATCCCTTGTCAATTATGACCTGCTGCCAGTCCGAGACAGTTCATCCTCCCCTAAGCAA 2159

Qy 3376 GAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGCAGTGAATGA 3423
 |||
 Db 2160 GAGGACAGTGGAGGTAGCAGCAGCAACAGCAGCAGCAGCAGTGAATGA 2207

RESULT 3

US-10-092-390-3

; Sequence 3, Application US/10092390

; Publication No. US20030013865A1

; GENERAL INFORMATION:

; APPLICANT: Yu, Xuanchuan

; APPLICANT: Miranda, Maricar

; TITLE OF INVENTION: No. US20030013865A1el Human EGF-Family Proteins and Polynucleotides Encoding the Same
 ; FILE REFERENCE: LEX-0317-USA
 ; CURRENT APPLICATION NUMBER: US/10/092,390
 ; CURRENT FILING DATE: 2002-03-06
 ; PRIOR APPLICATION NUMBER: US 60/275,013
 ; PRIOR FILING DATE: 2001-03-12
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1761
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 US-10-092-390-3

Query Match 51.4%; Score 1760; DB 14; Length 1761;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT	60
Db	1	ATGGTTATTTCTTTGAACTCATGCCTGAGCTTTATTTGTTTATTGTTATGCCACTGGATT	60
Qy	61	GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC	120
Db	61	GGGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTAC	120
Qy	121	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	180
Db	121	TCAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC	180
Qy	181	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	240
Db	181	ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTAT	240
Qy	241	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	300
Db	241	CGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAA	300
Qy	301	AGCGGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	360
Db	301	AGCGGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCT	360
Qy	361	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	420
Db	361	CCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCGAT	420
Qy	421	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	480
Db	421	GGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGC	480
Qy	481	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	540
Db	481	AACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGAC	540
Qy	541	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	600

Db	541	CGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGA	600
Qy	601	GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	660
Db	601	GCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCTTC	660
Qy	661	TGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Db	661	TGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGT	720
Qy	721	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Db	721	CAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATG	780
Qy	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Db	781	GGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAA	840
Qy	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Db	841	TGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCA	900
Qy	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960
Db	901	GGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGT	960
Qy	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1020
Db	961	GCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGGCGCATGCCTC	1020
Qy	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTAC	1080
Db	1021	TGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTAC	1080
Qy	1081	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTCACCCC	1140
Db	1081	GGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTCACCCC	1140
Qy	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Db	1141	ATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGT	1200
Qy	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Db	1201	TCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGAC	1260
Qy	1261	TGTGACAGTGTGACTGGAAAGTGACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Db	1261	TGTGACAGTGTGACTGGAAAGTGACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCT	1320
Qy	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAAT	1380
Db	1321	ACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAAT	1380
Qy	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGACTTGCAAGGCAGGCTGGCACGGGGTG	1440
Db	1381	GATGCAGTCTGCTCTCCTGTGGACGGGTCTTGACTTGCAAGGCAGGCTGGCACGGGGTG	1440

Qy 1441 GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACTTAACATGCCAG 1500
 |||
 Db 1441 GACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACTTAACATGCCAG 1500

Qy 1501 TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG 1560
 |||
 Db 1501 TGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGG 1560

Qy 1561 CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG 1620
 |||
 Db 1561 CGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAG 1620

Qy 1621 CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC 1680
 |||
 Db 1621 CGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTC 1680

Qy 1681 CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC 1740
 |||
 Db 1681 CCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAAC 1740

Qy 1741 TGCTCCCTGCCCTGCTACTG 1760
 |||
 Db 1741 TGCTCCCTGCCCTGCTACTG 1760

RESULT 4

US-10-105-929-9

; Sequence 9, Application US/10105929

; Publication No. US20020137142A1

; GENERAL INFORMATION:

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Goodearl, Andrew D.J.

; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83

; FILE REFERENCE: 09404/041001

; CURRENT APPLICATION NUMBER: US/10/105,929

; CURRENT FILING DATE: 2002-03-25

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108

; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961

; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 9

; LENGTH: 1448

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-105-929-9

Query Match 41.7%; Score 1425.8; DB 13; Length 1448;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1427; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1216 GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT 1275
 |||

Db 18 GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT 77

Qy 1276 GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG 1335
 |||

Db 78 GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG 137

Qy 1336 GGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT 1395
 |||

Db 138 GGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT 197

Qy 1396 CCTGTGGACGGGTCTTGTAAGGACAGGCTGGCACGGGGTGGACTGCTCCATCAGA 1455
 |||

Db 198 CCTGTGGACGGGTCTTGTAAGGACAGGCTGGCACGGGGTGGACTGCTCCATCAGA 257

Qy 1456 TGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTTAACATGCCAGTGCCTCAACGGGGGA 1515
 |||

Db 258 TGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTTAACATGCCAGTGCCTCAACGGGGGA 317

Qy 1516 GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC 1575
 |||

Db 318 GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC 377

Qy 1576 GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC 1635
 |||

Db 378 GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC 437

Qy 1636 CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT 1695
 |||

Db 438 CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT 497

Qy 1696 GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC 1755
 |||

Db 498 GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC 557

Qy 1756 TACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC 1815
 |||

Db 558 TACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC 617

Qy 1816 TTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTATGGGCATCGCTGCAGC 1875
 |||

Db 618 TTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTATGGGCATCGCTGCAGC 677

Qy 1876 CAGACATGCCCACAGTGCCTTACAGCAGCGGGCCCTGCCACCACATCACGGGCCTGTGT 1935
 |||

Db 678 CAGACATGCCCACAGTGCCTTACAGCAGCGGGCCCTGCCACCACATCACGGGCCTGTGT 737

Qy 1936 GACTGCTTGCCTGGCTTACAGGCGCCCTCTGCAATGAAGTGTGTCCAGTGGCAGATTT 1995
 |||

Db 738 GACTGCTTGCCTGGCTTACAGGCGCCCTCTGCAATGAAGTGTGTCCAGTGGCAGATTT 797

Qy 1996 GGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTTGAC 2055
 |||

Db 798 GGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTTGAC 857

Qy 2056 AGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTACTGCTCTCAACCATGTCCACCT 2115
 |||

Db 858 AGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTACTGCTCTCAACCATGTCCACCT 917

QY 2116 GCCCACTGGGGCCCAAACCTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC 2175
 |||
 Db 918 GCCCACTGGGGCCCAAACCTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC 977
 QY 2176 GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA 2235
 |||
 Db 978 GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA 1037
 QY 2236 TGTCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACGGAGCT 2295
 |||
 Db 1038 TGTCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACGGAGCT 1097
 QY 2296 GACTGCGACCACATTTCTGGGCAGTGACTTGCCGCACTGGATTTCATGGGACGGCACTGT 2355
 |||
 Db 1098 GACTGCGACCACATTTCTGGGCAGTGACTTGCCGCACTGGATTTCATGGGACGGCACTGT 1157
 QY 2356 GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGTCTG 2415
 |||
 Db 1158 GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGTCTG 1217
 QY 2416 AACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA 2475
 |||
 Db 1218 AACAACTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA 1277
 QY 2476 GCGAGATGTGATCAAGCTGGTGTTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC 2535
 |||
 Db 1278 GCGAGATGTGATCAAGCTGGTGTTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC 1337
 QY 2536 AGTACTGCTCTCCCTGCTGATTCCTACCAGATCGGGGCCATTGCAGGCATCATCATTCTT 2595
 |||
 Db 1338 AGTACTGCTCTCCCTGCTGATTCCTACCAATCGGGGCCATTGCAGGCATCATCATTCTT 1397
 QY 2596 GTCCTAGTTGTTCTCTTCCTACTGGCATTGTTTCATTATTTATAGACACA 2644
 |||
 Db 1398 GTCCTAGTTGTTCTCTTCCTACTGGCATTGTTTCATTATTTATAGACACA 1446

RESULT 5

US-10-365-227-9

; Sequence 9, Application US/10365227

; Publication No. US20030143632A1

; GENERAL INFORMATION:

; APPLICANT: McCarthy, Sean A.

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Goodearl, Andrew D.J.

; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING

; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND
 OTHER

; TITLE OF INVENTION: USES

; FILE REFERENCE: 07334-323001

; CURRENT APPLICATION NUMBER: US/10/365,227

; CURRENT FILING DATE: 2003-02-12

; PRIOR APPLICATION NUMBER: US/09/802,582

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: US 09/128,709

; PRIOR FILING DATE: 1998-08-04

```

; PRIOR APPLICATION NUMBER: US 60/054,645
; PRIOR FILING DATE: 1997-08-04
; PRIOR APPLICATION NUMBER: US 09/130,491
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: US 60/054,966
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 09/388,280
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 09/388,279
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1448
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-365-227-9

```

```

Query Match          41.7%; Score 1425.8; DB 14; Length 1448;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1427; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1216 GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT 1275
          |||
Db       18 GGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACT 77

Qy      1276 GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG 1335
          |||
Db       78 GGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTG 137

Qy      1336 GGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT 1395
          |||
Db       138 GGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCT 197

Qy      1396 CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGA 1455
          |||
Db       198 CCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGA 257

Qy      1456 TGTCCCACTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGA 1515
          |||
Db       258 TGTCCCACTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGCCTCAACGGGGGA 317

Qy      1516 GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC 1575
          |||
Db       318 GCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGC 377

Qy      1576 GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC 1635
          |||
Db       378 GAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGC 437

Qy      1636 CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT 1695
          |||
Db       438 CACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGT 497

Qy      1696 GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC 1755

```

Db	498	 GTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGC	557
Qy	1756	TACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC	1815
Db	558	 TACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGC	617
Qy	1816	TTCCGAGGCACCACTTGTCTCAGAGGATCTGCTCCCCGGTTTTTATGGGCATCGCTGCAGC	1875
Db	618	 TTCCGAGGCACCACTTGTCTCAGAGGATCTGCTCCCCGGTTTTTATGGGCATCGCTGCAGC	677
Qy	1876	CAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGT	1935
Db	678	 CAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGT	737
Qy	1936	GACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT	1995
Db	738	 GACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTT	797
Qy	1996	GGGAAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTTGAC	2055
Db	798	 GGGAAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCCCATTTGAC	857
Qy	2056	AGATCTTGTCTAGTGTACCCCGTTGGATTGGCAGTGACTGCTCTCAACCATGTCCACCT	2115
Db	858	 AGATCTTGTCTAGTGTACCCCGTTGGATTGGCAGTGACTGCTCTCAACCATGTCCACCT	917
Qy	2116	GCCCACTGGGGCCCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC	2175
Db	918	 GCCCACTGGGGCCCCAACTGCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGC	977
Qy	2176	GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA	2235
Db	978	 GCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGA	1037
Qy	2236	TGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCT	2295
Db	1038	 TGTCCTCTAGGGTTTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAACGGAGCT	1097
Qy	2296	GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGCACTGT	2355
Db	1098	 GACTGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTTCATGGGACGGCACTGT	1157
Qy	2356	GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGTCTG	2415
Db	1158	 GAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCGCCAGATATGTGATTGTCTG	1217
Qy	2416	AACAACCTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA	2475
Db	1218	 AACAACCTCCACCTGCGACCACATCACTGGGACCTGTTACTGCAGCCCCGGATGGAAGGGA	1277
Qy	2476	GCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC	2535
Db	1278	 GCGAGATGTGATCAAGCTGGTGTATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACC	1337
Qy	2536	AGTACTGCTCTCCCTGCTGATTCTTACCAGATCGGGGCCATTGCAGGCATCATCATTCTT	2595

Db 1338 AGTACTGCTCTCCCTGCTGATTCCCTACCAAATCGGGGCCATTGCAGGCATCATCATTCTT 1397

QY 2596 GTCCTAGTTGTTCTCTTCTACTGGCATTGTTTCATTATTTATAGACACA 2644
 |||

Db 1398 GTCCTAGTTGTTCTCTTCTACTGGCATTGTTTCATTATTTATAGACACA 1446

RESULT 6

US-10-052-648A-9

; Sequence 9, Application US/10052648A

; Publication No. US20040005558A1

; GENERAL INFORMATION:

; APPLICANT: Anderson, David

; APPLICANT: Burgess, Catherine

; APPLICANT: Casman, Stacie

; APPLICANT: Colman, Steven

; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Ellerman, Karen

; APPLICANT: Gerlach, Valerie

; APPLICANT: Gunther, Erik

; APPLICANT: Kekuda, Ramesh

; APPLICANT: MacDougall, John R.

; APPLICANT: Mehraban, Fuad

; APPLICANT: Patturajan, Meera

; APPLICANT: Rothenberg, Mark

; APPLICANT: Shimkets, Richard

; APPLICANT: Smithson, Glennda

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Stone, David J.

; APPLICANT: Vernet, Corine A.M.

; APPLICANT: Zerhusen, Bryan D.

; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF

; TITLE OF INVENTION: USING THE SAME

; FILE REFERENCE: 21402-250 (CURA-550)

; CURRENT APPLICATION NUMBER: US/10/052,648A

; CURRENT FILING DATE: 2002-12-09

; PRIOR APPLICATION NUMBER: 60/262,454

; PRIOR FILING DATE: 2001-01-18

; PRIOR APPLICATION NUMBER: 60/272,920

; PRIOR FILING DATE: 2001-03-02

; PRIOR APPLICATION NUMBER: 60/284,549

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/303,229

; PRIOR FILING DATE: 2001-07-05

; PRIOR APPLICATION NUMBER: 60/262,892

; PRIOR FILING DATE: 2001-01-19

; PRIOR APPLICATION NUMBER: 60/263,605

; PRIOR FILING DATE: 2001-01-23

; PRIOR APPLICATION NUMBER: 60/269,098

; PRIOR FILING DATE: 2001-02-15

; PRIOR APPLICATION NUMBER: 60/264,159

; PRIOR FILING DATE: 2001-01-25

; PRIOR APPLICATION NUMBER: 60/265,517

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/271,855

; PRIOR FILING DATE: 2001-02-27

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-052-648A-9

Query Match 18.7%; Score 640.6; DB 15; Length 3114;
Best Local Similarity 57.5%; Pred. No. 1.6e-196;
Matches 1235; Conservative 0; Mismatches 899; Indels 15; Gaps 4;

```
Qy      62 GGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACT 121
      ||  ||  | |||| || |  || || ||  || ||| ||||| ||||| |
Db      47 GGCTGGCTGGAACCTCTCAACCCAGTGATCCCAATACCTGCAGCTTCTGGGAAAGCTTCA 106

Qy     122 CAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC- 180
      |  ||  | |||| || | | |||| |  |  ||  | |||| 
Db     107 CTACCACCACCAAGGAGTCCCACTCCCGCCCTTCAGCCTGCTCCCCTCAGAGCCCTGCG 166

Qy     181 --ACTGACATTCTAAACTGGTTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCT 238
      |  |  |  || ||| | | || | |  || ||||| | | 
Db     167 AGCGGCCCTGGGAGGGCCCCCATACTTGCCCCCAGCCACGGTTGTATACCGGACCGTGT 226

Qy     239 ATCGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTGTCCCTGGATTTTATG 298
      | || | | || |||||  | | |  |||| | | | ||| | |||| 
Db     227 ACCGTCAGGTGGTGAAGACGGACCACCGCCAGCGCCTGCAGTGTGCCATGGCTTCTATG 286

Qy     299 AAAGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTG 358
      | ||| |||| | ||||| | |||| | | ||||| ||| ||| | | 
Db     287 AGAGCAGGGAGTTCTGTGTCCCGCTCTGTGCCCAGGAGTGTGTCCATGGCCGTTGTGTGG 346

Qy     359 CTCCAAACACCTGTGAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCG 418
      | || | |  || || |||| || ||||| | ||  | |||| ||||| || | 
Db     347 CACCCAATCAGTGCCAATGTGTGCCAGGCTGGCGGGGCGACGACTGTTCCAGTGAGTGTG 406

Qy     419 ATGGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGT 478
      |  |||| | || || | || | | || | || | ||  || 
Db     407 CCCCAGGAATGTGGGGGCCACAGTGTGACAAGCCCTGCAGTTGCGGCAACAACAGCTCGT 466

Qy     479 GCAACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGG 538
      | | |||| | |||| ||  || || | | | | | | |||| 
Db     467 GTGATCCCAAGAGTGGGGTATGTTCTTGCCCTTCTGGTCTGCAGCCCCGAACTGCCTTC 526

Qy     539 ACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATG 598
      | | |||| | ||| ||||| | || | |  | ||||| |||| | || 
Db     527 AGCCCTGTACCCCTGGCTACTATGGCCCTGCCTGCCAGTTCGCTGCCAGTGCC---ATG 583

Qy     599 GAGCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCT 658
      | ||  ||||| | |  || || | || | |||| | ||  || || || 
Db     584 GGGCACCTGCGATCCCCAGACTGGAGCCTGCTTCTGCCCCGAGAGAGAACTGGGCCCA 643

Qy     659 TCTGTGAGGATCTTTGTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTT 718
      ||||| |  ||| | | || | |||  ||  |  |||| 
Db     644 GCTGTGACGTGTCTGTTCAGGGCACTTCTGGCTTCTTCTGCCCCAGCACCCATCCTT 703

Qy     719 GTCAAAATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGA 778
```

Db	704	GCCAAAATGGAGGTGTCTTCCAAACCCACAGGGCTCCTGCAGCTGCCCCCTGGCTGGA	763
Qy	779	TGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAG	838
Db	764	TGGGCACCATCTGCTCCCTGCCCTGCCAGAGGGCTTTACGGACCCAACTGCTCCCAGG	823
Qy	839	AATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTC	898
Db	824	AATGTCGCTGCCACAACGGCGGCCTCTGTGACCGATTCACTGGGCAGTGCCGCTGCGCTC	883
Qy	899	CAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCT	958
Db	884	CGGGTTACACTGGGGATCGGTGCCGGGAGGAGTGCCCGGTGGGCCGCTTTGGGCAGGACT	943
Qy	959	GTGCTGAGACCTGCCAGTGTGTCAACGGAGGGGAAGTGTTACCACGTGAGCGGCGCATGCC	1018
Db	944	GTGCTGAGACGTGCGACTGCGCCCCGACGCCCCGTTGCTTCCCGGCCAACGGCGCATGTC	1003
Qy	1019	TCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCT	1078
Db	1004	TGTGCGAACACGGCTTCACTGGGGACCGCTGCACGGATCGCCTCTGCCCCGACGGCTTCT	1063
Qy	1079	ACGGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTCACC	1138
Db	1064	ACGGTCTCAGCTGCCAGGCCCCCTGCACCTGCGACCGGGAGCACAGCCTCAGCTGCCACC	1123
Qy	1139	CCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACAT	1198
Db	1124	CGATGAACGGGGAGTGCTCCTGCCTGCCGGGCTGGGCGGGCCTCCACTGCAACGAGAGCT	1183
Qy	1199	GTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAG	1258
Db	1184	GCCCGCAGGACACGCATGGGCCAGGGTGCCAGGAGTACTGTCTCTGCCTGCACGGTGGCG	1243
Qy	1259	ACTGTGACAGTGTGACTGGAAAGTGACCTGTGCCCCAGGATTCAAAGGAATTGACTGCT	1318
Db	1244	TCTGCCAGGCTACCAGCGGCCTCTGTCACTGCGCGCCGGGTTACAGGGCCCTCACTGTG	1303
Qy	1319	CTACCCCATGCCCTCTGGAACCTATGGGATAAACTGTTTCTCTCGCTGTGGCTGTAAAA	1378
Db	1304	CTAGTCTTTGTCCTCCTGACACCTACGGTGTCAACTGTTCTGCACGCTGCTCATGTGAAA	1363
Qy	1379	ATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGG	1438
Db	1364	ATGCCATCGCCTGCTCACCCATCGACGGCGAGTGCGTCTGCAAGGAAGGTTGGCAGCGTG	1423
Qy	1439	TGGACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTTAACATGCC	1498
Db	1424	GTAACCTGCTCTGTGCCCTGCCACCCGGAACCTGGGGCTTCAGTTGCAATGCCAGCTGCC	1483
Qy	1499	AGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGAT	1558
Db	1484	AGTGTGCCCATGAGGCAGTCTGCAGCCCCAACTGGAGCCTGTACCTGCACCCCTGGGT	1543
Qy	1559	GGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTG	1618

Db 1544 GGCATGGGGCCCACTGCCAGCTGCCCTGTCCGAAGGGGCAGTTTGGAGAAGGTTGTGCCA 1603
 Qy 1619 AGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGGCCATTGCCGCTGCC 1678
 |||| |||| |||| | ||||| |||| | | | | ||||
 Db 1604 GTCGCTGTGACTGTGACCACTCTGATGGCTGTGACCCTGTTCATGGACGCTGTCAGTGCC 1663
 Qy 1679 TCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCA 1738
 | | | | |||| | | | | | | | | |||| | |
 Db 1664 AGGCTGGCTGGATGGGTGCCCCTGCCACCTGTCCTGCCCTGAGGGCTTATGGGGAGTCA 1723
 Qy 1739 ACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCT 1798
 |||| | |||| | | | | |||| | | |||| |||| | |
 Db 1724 ACTGTAGCAACACCTGCACCTGCAAGAATGGGGGCACCTGTCTCCCTGAGAATGGCAACT 1783
 Qy 1799 GCGAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGACAGGATCTGCTCCCCTGGTTTTT 1858
 || | |||| || |||| | | | | | | | | |||| | |
 Db 1784 GCGTGTGTGCACCCGGATTCCGGGGCCCCCTCCTGCCAGAGATCCTGTCAGCCTGGCCGCT 1843
 Qy 1859 ATGGGCATCGCTGCAGCCAGACATGCCCACAGTGCCTTACAGCAGCGGGCCCTGCCACC 1918
 |||| | |||| | | | | |||| | | | | |||||
 Db 1844 ATGGCAAACGCTGTGTGC-----CCTGCAAGTGCCTAACCCTCC---TTCTGCCACC 1894
 Qy 1919 ACATCACCGGCCTGTGTGACTGCTTGCCTGGCTTACAGGCGCCCTCTGCAATGAAGTGT 1978
 | | | | || |||| | | |||| |||| | | | | | |
 Db 1895 CCTCGAACGGGACCTGCTACTGCCTGGCTGGCTGGACAGGCCCCGACTGCTCCCAGCCAT 1954
 Qy 1979 GTCCCACTGGCAGATTTGGGAAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAA 2038
 | | | | | | | | ||||| | | | | | | | | |
 Db 1955 GCCCTCCAGGACACTGGGGAGAAAACTGTGCCAGACCTGCCAATGTCACCATGGTGGGA 2014
 Qy 2039 CCTGTAACCCCATTTGACAGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGACTGCT 2098
 |||| | || | | | || | | | | | | | | |||||
 Db 2015 CCTGCCATCCCCAGGATGGGAGCTGTATCTGCCCCCTAGGCTGGACTGGACACCACTGCT 2074
 Qy 2099 CTCAACCATGTCCACCTGCCCCTGGGGCCCAACTGCATCCACACGTGCAACTGCCATA 2158
 | | | | | | | | | | | | | | | | | | | |
 Db 2075 TAGAAGGCTGCCCTCTGGGGACATTTGGTGCTAACTGCTCCCAGCCATGCCAGTGTGGTC 2134
 Qy 2159 ATGGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGG 2207
 |||| | | | | |||| | | | | | |
 Db 2135 CTGGAGAAAAGTGCCACCCAGAGACTGGGGCCTGTGTATGTCCCCCAGG 2183

RESULT 7

US-10-052-648A-7

; Sequence 7, Application US/10052648A

; Publication No. US20040005558A1

; GENERAL INFORMATION:

; APPLICANT: Anderson, David

; APPLICANT: Burgess, Catherine

; APPLICANT: Casman, Stacie

; APPLICANT: Colman, Steven

; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Ellerman, Karen

; APPLICANT: Gerlach, Valerie

; APPLICANT: Gunther, Erik

; APPLICANT: Kekuda, Ramesh

Db	249	AGCGGCCCTGGGAGGGCCCCATACTTGCCCCAGCCCACGGTTGTATACCGGACCGTGT	308
Qy	239	ATCGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATG	298
Db	309	ACCGTCAGGTGGTGAAGACGGACCACGCCAGCGCCTGCAGTGCTGCCATGGCTTCTATG	368
Qy	299	AAAGCGGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTG	358
Db	369	AGAGCAGGGGGTTCTGTGTCCCGCTCTGTGCCCAGGAGTGTGTCCATGGCCGTTGTGTGG	428
Qy	359	CTCCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCG	418
Db	429	CACCCAATCAGTGCCAATGTGTGCCAGGCTGGCGGGGCGACGACTGTTCCAGTGAGTGTG	488
Qy	419	ATGGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGT	478
Db	489	CCCCAGGAATGTGGGGGCCACAGTGTGACAAGCCCTGCAGCTGCGGCAACAACAGCTCGT	548
Qy	479	GCAACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGG	538
Db	549	GTGATCCCAAGAGTGGGGTATGTTCTTGCCCTTCTGGTCTGCAGCCCCCGAAGTGCCTTC	608
Qy	539	ACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATG	598
Db	609	AGCCCTGTACCCCTGGCTACTATGGCCCTGCCTGCCAGTTCCGCTGCCAGTGCC---ATG	665
Qy	599	GAGCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCT	658
Db	666	GGGCACCCTGCGATCCCCAGACTGGAGCCTGCTTCTGCCCCGCAGAGAGAACTGGGCCCA	725
Qy	659	TCTGTGAGGATCTTTGTCTCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTT	718
Db	726	GCTGTGACGTGTCTGTTCACAGGGCACTTCTGGCTTCTTCTGCCCCAGCACCCATCCTT	785
Qy	719	GTCAAATGGAGGAGTGTGTCACTGAGAGTGTCTCTTGCCCTTCTGGCTGGA	778
Db	786	GCCAAAATGGAGGTGCTTCCAAACCCACAGGGCTCCTGCAGCTGCCCCCTGGCTGGA	845
Qy	779	TGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAG	838
Db	846	TGGGCACCATCTGCTCCCTGCCCTGCCAGAGGGCTTTCACGGACCAACTGCTCCCAGG	905
Qy	839	AATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTC	898
Db	906	AATGTCGCTGCCACAACGGCGGCCTCTGTGACCGATTCACTGGGCAGTGCCGCTGCGCTC	965
Qy	899	CAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCT	958
Db	966	CGGGTTACACTGGGGATCGGTGCCGGGAGGAGTGCCCGGTGGGCGCTTTGGGCAGGACT	1025
Qy	959	GTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACACGTGAGCGGCGCATGCC	1018
Db	1026	GTGCTGAGACGTGCGACTGCGCCCCGACGCCCCTTGCTTCCCGGCCAACGGCGCATGTC	1085
Qy	1019	TCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCCTGAGGGGCTCT	1078

Db 1086 TGTGCGAACACGGCTTCACTGGGGACCGCTGCACGGATCGCCTCTGCCCCGACGGCTTCT 1145
 Qy 1079 ACGGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTCACC 1138
 |||| ||| || | | | |||| || ||| ||| | |||| ||||
 Db 1146 ACGGTCTCAGCTGCCAGGCCCGCCGACCTGCGACCGGGAGCACAGCCTCAGCTGCCACC 1205
 Qy 1139 CCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACAT 1198
 | ||| || ||||| ||||| ||||| || | || ||| |||| || |||| |
 Db 1206 CGATGAACGGGGAGTGCTCCTGCCTGCCGGGCTGGGCGGGCCTCCACTGCAACGAGAGCT 1265
 Qy 1199 GTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAG 1258
 | | | | | ||| || ||||| || ||| |||| | || | |
 Db 1266 GCGCGCAGGACACGCATGGGCCAGGGTGCCAGGAGCACTGTCTCTGCCTGCACGGTGGCG 1325
 Qy 1259 ACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCT 1318
 ||| | | | || || || || || || || || || || || || ||
 Db 1326 TCTGCCAGGCTACCAGCGGCCTCTGTCAGTGCGCGCCGGGTACACGGGCCCTCACTGTG 1385
 Qy 1319 CTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCTCTCGCTGTGGCTGTAAAA 1378
 ||| | || |||| | |||| || | ||||| || |||| || || ||
 Db 1386 CTAGTCTTTGTCTCCTGACACCTACGGTGTCACTGTTCTGCACGCTGCTCATGTGAAA 1445
 Qy 1379 ATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGG 1438
 ||| | ||||| || | |||| || || ||||| || |||| | |
 Db 1446 ATGCCATCGCCTGCTCACCCATCGACGGCGAGTGCCTGCAAGGAAGGTTGGCAGCGTG 1505
 Qy 1439 TGGACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCC 1498
 ||||| | || || || || || ||||| || || || | ||||
 Db 1506 GTAACCTGCTCTGTGCCCTGCCCACCGGAACCTGGGGCTTCAGTTGCAATGCCAGCTGCC 1565
 Qy 1499 AGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGAT 1558
 |||| | | || || || |||| || || || || || || || || || ||
 Db 1566 AGTGTGCCCATGAGGCAGTCTGCAGCCCCCAAACCTGGAGCCTGTACCTGCACCCCTGGGT 1625
 Qy 1559 GGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTG 1618
 ||| |||| | ||| || || |||| || || || || || || || || ||
 Db 1626 GGATGGGGCCCACTGCCAGCTGCCCTGTCCGAAGGGGCAGTTTGAGAGAAGGTTGTGCCA 1685
 Qy 1619 AGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCC 1678
 |||| |||| |||| | ||||| |||| || || || || || || || ||
 Db 1686 GTCGCTGTGACTGTGACCACTCTGATGGCTGTGACCCTGTTTCATGGACGCTGTCAGTGCC 1745
 Qy 1679 TCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCA 1738
 | || ||| |||| || || || || || || || || || || || || ||
 Db 1746 AGGCTGGCTGGATGGGTGCCCCGCTGCCACCTGTCCTGCCCTGAGGGCTTATGGGGAGTCA 1805
 Qy 1739 ACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCT 1798
 |||| | |||| ||| || ||||| || || ||||| ||||| || ||
 Db 1806 ACTGTAGCAACACCTGCACCTGCAAGAATGGGGGCACCTGTCTCCCTGAGAATGGCAACT 1865
 Qy 1799 GCGAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTT 1858
 ||| ||||| || || ||||| ||| || || ||||| ||| ||||| |
 Db 1866 GCGTGTGTGCGCCCGGATTCCGGGGCCCCCTCCTGCCAGAGATCCTGTCAGCCTGGCCGCT 1925
 Qy 1859 ATGGGCATCGCTGCAGCCAGACATGCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACC 1918
 |||| | |||| | | | |||| | || | | ||||| |||||
 Db 1926 ATGGCAAACGCTGTGTGC-----CCTGCAAGTGCCTAACCCTCC---TTCTGCCACC 1976

Qy 1919 ACATCACC GGCTGTGTGACTGCTTGCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGT 1978
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1977 CCTCGAACGGGACCTGCTACTGCCTGGCTGGCTGGACAGGCCCCGACTGCTCCCAGCCAT 2036
 Qy 1979 GTCCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAA 2038
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2037 GCCCTCCAGGACACTGGGGAGAAAACTGTGCCAGACCTGCCAATGTCACCATGGTGGGA 2096
 Qy 2039 CCTGTAACCCCATTTGACAGATCTTGTCAAGTGTACCCCGGTTGGATTGGCAGTGAAGTGT 2098
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2097 CCTGCCATCCCCAGGATGGGAGCTGTATCTGCCCCCTAGGCTGGACTGGACACCACTGCT 2156
 Qy 2099 CTCAACCATGTCCACCTGCCACTGGGGCCCAAACCTGCATCCACACGTGCAACTGCCATA 2158
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2157 TAGAAGGCTGCCCTCTGGGGACATTTGGTGCTAACTGCTCCCAGCCATGCCAGTGTGGTC 2216
 Qy 2159 ATGGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGG 2207
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2217 CTGGAGAAAAGTGCCACCCAGAGACTGGGGCCTGTGTATGTCCCCCAGG 2265

RESULT 8

US-09-796-753-99

; Sequence 99, Application US/09796753
 ; Publication No. US20030027998A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCarthy, Sean A.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
 ; FILE REFERENCE: 7853-227-999
 ; CURRENT APPLICATION NUMBER: US/09/796,753
 ; CURRENT FILING DATE: 2001-03-01
 ; PRIOR APPLICATION NUMBER: 09/183,175
 ; PRIOR FILING DATE: 1998-10-30
 ; PRIOR APPLICATION NUMBER: 09/223,094
 ; PRIOR FILING DATE: 1998-12-30
 ; PRIOR APPLICATION NUMBER: 09/223,546
 ; PRIOR FILING DATE: 1998-12-30
 ; PRIOR APPLICATION NUMBER: 09/224,246
 ; PRIOR FILING DATE: 1998-12-30
 ; PRIOR APPLICATION NUMBER: 09/259,388
 ; PRIOR FILING DATE: 1999-02-26
 ; PRIOR APPLICATION NUMBER: 60/122,458
 ; PRIOR FILING DATE: 1999-03-01
 ; PRIOR APPLICATION NUMBER: 09/312,359
 ; PRIOR FILING DATE: 1999-05-14
 ; PRIOR APPLICATION NUMBER: 09/336,536
 ; PRIOR FILING DATE: 1999-06-18
 ; PRIOR APPLICATION NUMBER: 09/342,687
 ; PRIOR FILING DATE: 1999-06-29
 ; PRIOR APPLICATION NUMBER: 09/345,464
 ; PRIOR FILING DATE: 1999-06-30
 ; PRIOR APPLICATION NUMBER: 09/365,164
 ; PRIOR FILING DATE: 1999-07-30
 ; PRIOR APPLICATION NUMBER: 09/399,723
 ; PRIOR FILING DATE: 1999-09-20
 ; PRIOR APPLICATION NUMBER: 09/409,634

```

; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 99
; LENGTH: 3567
; TYPE: DNA
; ORGANISM: Rauttus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (925)...(2832)
US-09-796-753-99

```

```

Query Match          17.8%; Score 608.4; DB 10; Length 3567;
Best Local Similarity 56.7%; Pred. No. 5.2e-186;
Matches 1226; Conservative 0; Mismatches 901; Indels 37; Gaps 4;

```

```

Qy      67 GCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACTCAGTG 126
      ||  |  ||  ||  |  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      236 GCTGGAACACTCAACTCCAATGATCCCAATGTCTGTACCTTCTGGGAAAGCTTACCACG 295

Qy      127 ACTGTGCAAGAGTCATACCCACATCCCTTT---GATCAAATTTACTACACGAGCTGCACT 183
      ||  |  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      296 ACCACTAAGGAGTCCCACCTTCGCCCTTCAGCCTGCCCCAGCCGAGTCCTGCGACAGG 355

Qy      184 GACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTATCGA 243
      |  |  ||  |  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      356 CCCTGGGAAGACCCCCACACCTGCGCTCAGCCTACGGTTGTCTACCGACTGTGTACCGT 415

Qy      244 CATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTGTCTCTGGATTTTATGAAAGC 303
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      416 CAGGTGGTGAAGATGGACTCCCGCCACGCCTGCAGTGCTGTGGGGGTTACTACGAGAGC 475

```

Qy 304 GGGGAAATGTGTGTCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCTCCA 363
| | | | | | | | | | | | | | | | | | | | | |
Db 476 AGTGGAGCCTGTGTCCCACCTCTGTGCCCAGGAGTGTTGTCCACGGTCGCTGTGTGGCTCCT 535

Qy 364 AACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCACACTGCTCCAGTGCCTGCGATGGT 423
|| | | | | | | | | | | | | | | | | | | | |
Db 536 AATCGGTGCCAGTGTGCACCAGGCTGGCGGGGTGACGACTGTTCCAGTGAGTGTGCTCCT 595

Qy 424 GATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGCAAC 483
| | | | | | | | | | | | | | | | | | | | |
Db 596 GGAATGTGGGGACCACAGTGTGACAGGCTCTGCCTCTGTGGCAACAGCAGTTCCTGTGAT 655

Qy 484 CCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGC 543
| | | | | | | | | | | | | | | | | | | | |
Db 656 CCCAGGAGTGGGGTGTGTTTTTGCCCTCTGGCCTGCAGCCCCCGACTGCCTTCAGCCT 715

Qy 544 TGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGAGCC 603
|| | | | | | | | | | | | | | | | | | | |
Db 716 TGCCCCGATGGCCACTATGGTCTGCCTGCCAGTTTGATTGCCATTGC---TATGGGGCA 772

Qy 604 ACCTGCGACCACGTACACGGGGGAATGCCGCTGCCCCACCAGGATACACCGGAGCCTTCTGT 663
| | | | | | | | | | | | | | | | | | | | |
Db 773 TCCTGTGACCCCCGGGATGGAGCCTGCTTCTGCCCCCAGGGAGAACAGGACCCAGGGCA 832

Qy 664 GAGGATCTTTGTCTCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAA 723
| | | | | | | | | | | | | | | | | | | | |
Db 833 CTGATGGCTTCTTCTGCCCCAGAAC-----TTATCCTTGCCAA 870

Qy 724 AATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGC 783
| | | | | | | | | | | | | | | | | | | | |
Db 871 AATGGAGGTGTTCTCAGGGCTCTCAAGGCTCCTGCAGCTGCCACCGGGCTGGATGGGT 930

Qy 784 ACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGC 843
| | | | | | | | | | | | | | | | | | | | |
Db 931 GTCATCTGTTCCCTGCCATGCCAGAGGGTTTCCACGGACCCAAGTGTACTCAGGAATGT 990

Qy 844 CAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGA 903
| | | | | | | | | | | | | | | | | | | | |
Db 991 CGTTGCCACAATGGTGGCCTTTGTGACAGGTTTACTGGGCAGTGCCACTGTGCTCCTGGC 1050

Qy 904 TACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCT 963
|| | | | | | | | | | | | | | | | | | | |
Db 1051 TATATCGGGGATCGGTGCCGTGAAGAGTGCCCTGTGGGCCGCTTCGGTCAAGACTGTGCT 1110

Qy 964 GAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTTACCACGTGAGCGGCGCATGCCTCTGT 1023
| | | | | | | | | | | | | | | | | | | | |
Db 1111 GAGACCTGTGACTGTGCTCCTGGCGCTCGTTGCTTTTCTGCCAATGGCGCGTGTCTGTGC 1170

Qy 1024 GAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTACGGC 1083
|| | | | | | | | | | | | | | | | | | | |
Db 1171 GAACATGGCTTCACAGGCGACCGCTGCACTGAGCGACTCTGTCCAGATGGCCGCTATGGT 1230

Qy 1084 ATCAAATGTGACAAACGGTGTCCCTGCCACTTGAAAAACACTCATAGCTGTACCCCCATG 1143
| | | | | | | | | | | | | | | | | | | | |
Db 1231 CTGAGCTGCCAAGATCCCTGCACCTGCGACCCAGAACACAGTCTCAGCTGCCACCCAATG 1290

Qy 1144 TCTGGAGAGTGTGCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCT 1203

Db	1291	<div> <div> </div> CACGGCGAGTGCTCCTGCCAGCCAGGTTGGGCGGGCCTCCACTGCAACGAGAGCTGCCCT </div>	1350
Qy	1204	CCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGT	1263
Db	1351	<div> <div> </div> CAGGACACGCACGGAGCCGGTTGCCAGGAGCACTGCCTCTGTCTGCACGGCGGTGTTTGC </div>	1410
Qy	1264	GACAGTGTGACTGGAAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACC	1323
Db	1411	<div> <div> </div> CTCGCCGACAGCGGCCTCTGCCGGTGTGCACCTGGCTACACGGGACCTCACTGCGCTAAT </div>	1470
Qy	1324	CCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAATGAT	1383
Db	1471	<div> <div> </div> CTTTGTCCACCTAACACTTATGGGATCAACTGTTCTCCCACTGCTCCTGTGAAAATGCC </div>	1530
Qy	1384	GCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCAGGGGTGGAC	1443
Db	1531	<div> <div> </div> ATTGCCCTGCTCTCCTGTGCACGGCACGTGCATCTGCAAGGAAGTTGGCAGCGTGGTAAC </div>	1590
Qy	1444	TGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGC	1503
Db	1591	<div> <div> </div> TGCTCTGTGCCCTGTCCCCCTGGCACCTGGGGCTTCAGTTGCAATGCCAGTTGCCAGTGT </div>	1650
Qy	1504	CTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGC	1563
Db	1651	<div> <div> </div> GCCCACGAGGGAGTCTGCAGCCCCCAAACCTGGAGCCTGTACTTGCACCCCTGGGTGGCGT </div>	1710
Qy	1564	GGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGC	1623
Db	1711	<div> <div> </div> GGGGTTCACTGCCAACTTCCGTGCCCCAAGGGACAGTTTGGTGAAGGTTGTGCCAGTGTC </div>	1770
Qy	1624	TGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCG	1683
Db	1771	<div> <div> </div> TGTGACTGTGACCACTCCGATGGCTGTGACCCTGTTTCATGGACACTGCCGATGTCAGGCT </div>	1830
Qy	1684	GGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGC	1743
Db	1831	<div> <div> </div> GGCTGGATGGGCACACGTTGCCACCTGCCTTGCCCAGAGGGCTTTTGGGGAGCCAACTGC </div>	1890
Qy	1744	TCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAG	1803
Db	1891	<div> <div> </div> AGCAATGCCTGTACCTGCAAGAATGGTGGCACTTGTGTACCTGAGAACGGCAACTGTGTG </div>	1950
Qy	1804	TGTGCACCAGGCTTCCGAGGCACCACTTGTGAGGATCTGCTCCCCTGGTTTTTATGGG	1863
Db	1951	<div> <div> </div> TGCGCACCAGGGTTCAGAGGCCCTCCTGCCAGAGGCCCTGCCCGCTGGTTCGCTATGGC </div>	2010
Qy	1864	CATCGCTGCAGCCAGACATGCCCACAGTGCCTTCACAGCAGCGGGCCCTGCCACCACATC	1923
Db	2011	<div> <div> </div> AAACGCTG-----TGTGCCCTGCAAGTGCAACAACCATTCTTCCTGCCACCCGTCG </div>	2061
Qy	1924	ACCGGCCTGTGTGACTGCTTGCCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCC	1983
Db	2062	<div> <div> </div> GATGGGACCTGCTCCTGCCTGGCAGGCTGGACAGGCCCTGACTGCTCTGAATCATGTCCC </div>	2121
Qy	1984	AGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGT	2043
		<div> <div> </div> </div>	

```

Db      2122 CCAGGCCACTGGGGACTCAAATGCTCCCAACCCTGCCAGTGTTCATCATGGTGCCACCTGC 2181
Qy      2044 AACCCCATTTGACAGATCTTGTCTAGTGTTACCCCGGTTGGATTGGCAGTGACTGCTCTCAA 2103
        |||||  ||  |   |||  ||  |||  ||  ||||  |||  |||||  ||
Db      2182 CACCCCCAGGATGGGAGCTGTGTCTGCATCCCAGGCTGGACTGGACCCAAGTCTCGGAA 2241
Qy      2104 CCATGTCCACCTGCCCCACTGGGGCCCAAAGTGCATCCACACGTGCAACTGCCATAATGGA 2163
        || ||| |   |  ||  |||||  |||  ||  |  ||  ||  ||||
Db      2242 GGCTGCCCATCAAGAATGTTTGGTGTCAACTGCTCCCAGCTATGTCAGTGTGATCCTGGA 2301
Qy      2164 GCTTTCTGCAGCGCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTAC 2223
        |  | |||  |  |  ||||  ||  ||  |  ||  ||  |  ||  ||
Db      2302 GAGATGTGCCACCCAGAGACTGGGGCTTGCGTCTGTCCCCCAGGACACAGTGGTGCGCAC 2361
Qy      2224 TGCA 2227
        ||||
Db      2362 TGCA 2365

```

RESULT 9

US-09-796-753-123

; Sequence 123, Application US/09796753

; Publication No. US20030027998A1

; GENERAL INFORMATION:

; APPLICANT: McCarthy, Sean A.

; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF

; FILE REFERENCE: 7853-227-999

; CURRENT APPLICATION NUMBER: US/09/796,753

; CURRENT FILING DATE: 2001-03-01

; PRIOR APPLICATION NUMBER: 09/183,175

; PRIOR FILING DATE: 1998-10-30

; PRIOR APPLICATION NUMBER: 09/223,094

; PRIOR FILING DATE: 1998-12-30

; PRIOR APPLICATION NUMBER: 09/223,546

; PRIOR FILING DATE: 1998-12-30

; PRIOR APPLICATION NUMBER: 09/224,246

; PRIOR FILING DATE: 1998-12-30

; PRIOR APPLICATION NUMBER: 09/259,388

; PRIOR FILING DATE: 1999-02-26

; PRIOR APPLICATION NUMBER: 60/122,458

; PRIOR FILING DATE: 1999-03-01

; PRIOR APPLICATION NUMBER: 09/312,359

; PRIOR FILING DATE: 1999-05-14

; PRIOR APPLICATION NUMBER: 09/336,536

; PRIOR FILING DATE: 1999-06-18

; PRIOR APPLICATION NUMBER: 09/342,687

; PRIOR FILING DATE: 1999-06-29

; PRIOR APPLICATION NUMBER: 09/345,464

; PRIOR FILING DATE: 1999-06-30

; PRIOR APPLICATION NUMBER: 09/365,164

; PRIOR FILING DATE: 1999-07-30

; PRIOR APPLICATION NUMBER: 09/399,723

; PRIOR FILING DATE: 1999-09-20

; PRIOR APPLICATION NUMBER: 09/409,634

; PRIOR FILING DATE: 1999-09-30

; PRIOR APPLICATION NUMBER: 09/471,179

; PRIOR FILING DATE: 1999-12-23


```

; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 123
; LENGTH: 3567
; TYPE: DNA
; ORGANISM: Rauttus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (925)...(2832)
US-09-796-753-123

```

```

Query Match          17.8%; Score 608.4; DB 10; Length 3567;
Best Local Similarity 56.7%; Pred. No. 5.2e-186;
Matches 1226; Conservative 0; Mismatches 901; Indels 37; Gaps 4;

```

```

Qy      67 GCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACTCAGTG 126
      ||  |  ||  ||  |  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      236 GCTGGAACACTCAACTCCAATGATCCCAATGTCTGTACCTTCTGGGAAAGCTTCACCACG 295

Qy      127 ACTGTGCAAGAGTCATACCCACATCCCTTT---GATCAAATTTACTACACGAGCTGCACT 183
      ||  |  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      296 ACCACTAAGGAGTCCCACCTTCGCCCCCTTCAGCCTGCCCCCAGCCGAGTCCTGCGACAGG 355

Qy      184 GACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCTATCGA 243
      |  |  |  ||  |  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      356 CCCTGGGAAGACCCCCACACCTGCGCTCAGCCTACGGTTGTCTACCGGACTGTGTACCGT 415

Qy      244 CATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCCTGGATTTTATGAAAGC 303
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      416 CAGGTGGTGAAGATGGACTCCCCGCCACGCCTGCAGTGCTGTGGGGTTACTACGAGAGC 475

Qy      304 GGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTGCTCCA 363
      |  |  |  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      476 AGTGGAGCCTGTGTCCCACTCTGTGCCCAGGAGTGTGTCCACGGTCGCTGTGTGGCTCCT 535

```

Qy	364	AACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCCTGCGATGGT	423
Db	536	AATCGGTGCCAGTGTGCACCAGGCTGGCGGGGTGACGACTGTTCCAGTGAGTGTGCTCCT	595
Qy	424	GATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGTGCAAC	483
Db	596	GGAATGTGGGGACCACAGTGTGACAGGCTCTGCCTCTGTGGCAACAGCAGTTCTCTGTGAT	655
Qy	484	CCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGGACCGC	543
Db	656	CCCAGGAGTGGGGTGTGTTTTTGGCCCTCTGGCCTGCAGCCCCCGACTGCCTTCAGCCT	715
Qy	544	TGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATGGAGCC	603
Db	716	TGCCCCGATGGCCACTATGGTCCTGCCTGCCAGTTTGATTGCCATTGC---TATGGGGCA	772
Qy	604	ACCTGCGACCACGTACACGGGGGAATGCCGCTGCCACCAGGATACACCGGAGCCTTCTGT	663
Db	773	TCCTGTGACCCCCGGGATGGAGCCTGCTTCTGCCCCCAGGGAGAACAGGACCCAGGGCA	832
Qy	664	GAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTTGTCAA	723
Db	833	CTGATGGCTTCTTCTGCCCCAGAAC-----TTATCCTTGCCAA	870
Qy	724	AATGGAGGAGTGTGTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGATGGGC	783
Db	871	AATGGAGGTGTTTCCTCAGGGCTCTCAAGGCTCCTGCAGCTGCCACCGGGCTGGATGGGT	930
Qy	784	ACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTGTTCCCAAGAATGC	843
Db	931	GTCATCTGTTCCCTGCCATGCCAGAGGGTTTCCACGGACCCAAGTGTACTCAGGAATGT	990
Qy	844	CAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCATTGCAGTCCAGGA	903
Db	991	CGTTGCCACAATGGTGGCCTTTGTGACAGGTTTACTGGGCAGTGCCACTGTGCTCCTGGC	1050
Qy	904	TACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCT	963
Db	1051	TATATCGGGGATCGGTGCCGTGAAGAGTGCCCTGTGGGCCGCTTCGGTCAAGACTGTGCT	1110
Qy	964	GAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTTACCACGTGAGCGGCGCATGCCCTCTGT	1023
Db	1111	GAGACCTGTGACTGTGCTCCTGGCGCTCGTTGCTTTCCTGCCAATGGCGCGTGTCTGTGC	1170
Qy	1024	GAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTTGAGGGGCTCTACGGC	1083
Db	1171	GAACATGGCTTCACAGGCGACCGCTGCACTGAGCGACTCTGTCCAGATGGCCGCTATGGT	1230
Qy	1084	ATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAAAACACTCATAGCTGTACCCCATG	1143
Db	1231	CTGAGCTGCCAAGATCCCTGCACCTGCGACCCAGAACACAGTCTCAGCTGCCACCCAATG	1290
Qy	1144	TCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCT	1203
Db	1291	CACGGCGAGTGCTCCTGCCAGCCAGGTTGGGCGGGCCTCCACTGCAACGAGAGCTGCCCT	1350

Qy 1204 CCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGT 1263
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1351 CAGGACACGCACGGAGCCGGTTGCCAGGAGCACTGCCTCTGTCTGCACGGCGGTGTTTGC 1410

Qy 1264 GACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACC 1323
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1411 CTCGCCGACACGGCCTCTGCCGGTGTGCACCTGGCTACACGGGACCTCACTGCGCTAAT 1470

Qy 1324 CCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGGCTGTAAAAATGAT 1383
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1471 CTTTGTCCACCTAACACTTATGGGATCAACTGTTCTCCCACTGCTCCTGTGAAAATGCC 1530

Qy 1384 GCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGAC 1443
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1531 ATTGCCTGCTCTCCTGTGACGGCACGTGCATCTGCAAGGAAGGTTGGCAGCGTGGTAAC 1590

Qy 1444 TGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCCAGTGC 1503
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1591 TGCTCTGTGCCCTGTCCCCCTGGCACCTGGGGCTTCAGTTGCAATGCCAGTTGCCAGTGT 1650

Qy 1504 CTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGCACCTGGATGGCGC 1563
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1651 GCCCACGAGGGAGTCTGCAGCCCCCAAACCTGGAGCCTGTACTTGCACCCCTGGGTGGCGT 1710

Qy 1564 GGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAACTGTGCTGAGCGC 1623
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1711 GGGGTTCACTGCCAACTTCCGTGCCCCGAAGGGACAGTTTGGTGAAGGTTGTGCCAGTGTG 1770

Qy 1624 TGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTGCCGCTGCCTCCCG 1683
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1771 TGTGACTGTGACCACTCCGATGGCTGTGACCCTGTTTCATGGACACTGCCGATGTCAGGCT 1830

Qy 1684 GGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGC 1743
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1831 GGCTGGATGGGCACACGTTGCCACCTGCCTTGCCAGAGGGCTTTTGGGGAGCCAACCTGC 1890

Qy 1744 TCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCTGCGAG 1803
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1891 AGCAATGCCTGTACCTGCAAGAATGGTGGCACTTGTGTACCTGAGAACGGCAACTGTGTG 1950

Qy 1804 TGTGCACCAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCCTGGTTTTTATGGG 1863
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1951 TGCGCACCAGGGTTCAGAGGCCCTCCTGCCAGAGGCCCTGCCCGCTGGTTCGCTATGGC 2010

Qy 1864 CATCGCTGCAGCCAGACATGCCACAGTGCCTTACAGCAGCGGGCCCTGCCACCACATC 1923
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2011 AAACGCTG-----TGTGCCCTGCAAGTGCAACAACCATTCTTCCTGCCACCCGTCG 2061

Qy 1924 ACCGGCCTGTGTGACTGCTTGCCCTGGCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCC 1983
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2062 GATGGGACCTGCTCCTGCCTGGCAGGCTGGACAGGCCCTGACTGCTCTGAATCATGTCCC 2121

Qy 1984 AGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGT 2043
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2122 CCAGGCCACTGGGGACTCAAATGCTCCCAACCCTGCCAGTGTGCATCATGGTGCCACCTGC 2181

Qy 2044 AACCCCATTGACAGATCTTGTGAGTGTACCCCGGTTGGATTGGCAGTGACTGCTCTCAA 2103

```

      |||||  || |  |||  ||  ||| || |||| |||  ||||| ||
Db      2182 CACCCCCAGGATGGGAGCTGTGTCTGCATCCCAGGCTGGACTGGACCCAAGTCTCGGAA 2241

Qy      2104 CCATGTCCACCTGCCCCTGGGGCCCAAGTGCATCCACACGTGCAACTGCCATAATGGA 2163
      || ||| |  | ||  ||||| |||  || | || || |||
Db      2242 GGCTGCCCATCAAGAATGTTTGGTGTCAACTGCTCCCAGCTATGTCAGTGTGATCCTGGA 2301

Qy      2164 GCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGGCTGGACAGGGCTCTAC 2223
      |  | ||| | | |  |||| |  || | || ||  | ||  ||
Db      2302 GAGATGTGCCACCCAGAGACTGGGGCTTGCGTCTGTCCCCCAGGACACAGTGGTGCAC 2361

Qy      2224 TGCA 2227
      ||||
Db      2362 TGCA 2365

```

RESULT 10

US-10-052-648A-3

; Sequence 3, Application US/10052648A

; Publication No. US20040005558A1

; GENERAL INFORMATION:

; APPLICANT: Anderson, David

; APPLICANT: Burgess, Catherine

; APPLICANT: Casman, Stacie

; APPLICANT: Colman, Steven

; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Ellerman, Karen

; APPLICANT: Gerlach, Valerie

; APPLICANT: Gunther, Erik

; APPLICANT: Kekuda, Ramesh

; APPLICANT: MacDougall, John R.

; APPLICANT: Mehraban, Fuad

; APPLICANT: Patturajan, Meera

; APPLICANT: Rothenberg, Mark

; APPLICANT: Shimkets, Richard

; APPLICANT: Smithson, Glennnda

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Stone, David J.

; APPLICANT: Vernet, Corine A.M.

; APPLICANT: Zerhusen, Bryan D.

; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF

; TITLE OF INVENTION: USING THE SAME

; FILE REFERENCE: 21402-250 (CURA-550)

; CURRENT APPLICATION NUMBER: US/10/052,648A

; CURRENT FILING DATE: 2002-12-09

; PRIOR APPLICATION NUMBER: 60/262,454

; PRIOR FILING DATE: 2001-01-18

; PRIOR APPLICATION NUMBER: 60/272,920

; PRIOR FILING DATE: 2001-03-02

; PRIOR APPLICATION NUMBER: 60/284,549

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/303,229

; PRIOR FILING DATE: 2001-07-05

; PRIOR APPLICATION NUMBER: 60/262,892

; PRIOR FILING DATE: 2001-01-19

; PRIOR APPLICATION NUMBER: 60/263,605

; PRIOR FILING DATE: 2001-01-23

```
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 2919
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-052-648A-3
```

Qy	62	GGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACT	121
Db	129	GGCTGGCTGGAACCTCAACCCCCAGTGATCCCAATACCTGCAGCTTCTGGGAAAGCTTCA	188
Qy	122	CAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC-	180
Db	189	CTACCACCACCAAGGAGTCCCCTCCCGCCCCCTTCAGCCTGCTCCCCTCAGAGCCCTGCG	248
Qy	181	--ACTGACATTTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCT	238
Db	249	AGCGGCCCTGGGAGGGCCCCCATACTTGCCCCCAGCCACGGTTGTATACCGGACCGTGT	308
Qy	239	ATCGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGT'TGTCCTGGATTTTATG	298
Db	309	ACCGTCAGGTGGTGAAGACGGACCACCGCCAGCGCCTGCAGTGCTGCCATGGCTTCTATG	368
Qy	299	AAAGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTG	358
Db	369	AGAGCAGGGGGTTCTGTGTCCCGCTCTGTGCCCAGGAGTGTGTCCATGGCCGTTGTGTGG	428
Qy	359	CTCCAAACACCTGTCTAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCG	418
Db	429	CACCCAATCAGTGCCAAATGTGTGCCAGGCTGGCGGGGCGACGACTGTTCCAGTGAGTGTG	488
Qy	419	ATGGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGT	478
Db	489	CCCCAGGAATGTGGGGGCCACAGTGTGACAAGCCCTGCAGCTGCGGCAACAACAGCTCGT	548
Qy	479	GCAACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGG	538
Db	549	GTGATCCCAAGAGTGGGGTATGTTCTTGCCCTTCTGGTCTGCAGCCCCGAAGTGCCTTC	608
Qy	539	ACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATG	598
Db	609	AGCCCTGTACCCCTGGCTACTATGGCCCTGCCTGCCAGTTCGCTGCCAGTGCC---ATG	665

Qy 599 GAGCCACCTGCGACCACGTACGCGGGGAATGCCGCTGCCACCCAGGATACACCGGAGCCT 658
 | | | | | | | | | | | | | | | | | | | | | |
 Db 666 GGGCACCTGCGATCCCCAGACTGGAGCCTGCTTCTGCCCCGAGAGAGAACTGGGCCCA 725

Qy 659 TCTGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTT 718
 | | | | | | | | | | | | | | | | | | | | | |
 Db 726 GCTGTGACGTGTCTGTTCACAGGGCACTTCTGGCTTCTTCTGCCCCAGCACCCATTCTT 785

Qy 719 GTCAAAATGGAGGAGTGTGTATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGA 778
 | | | | | | | | | | | | | | | | | | | | | |
 Db 786 GCCAAAATGGAGGTGTCTTCCAAACCCACAGGGCTCCTGCAGCTGCCCCCTGGCTGGA 845

Qy 779 T----- 779
 |
 Db 846 TGGTATGGAGGTGGGGCCTGTGGGCATGGGGTGTGGGTCTGGGGAGAATTCTGTGGGTG 905

Qy 780 -----GGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTG 820
 | | | | | | | | | | | | | | | | | | | | | |
 Db 906 GTGCTAAGCAGGGCTCCAAGGGCACCATCTGCTCCCTGCCCTGCCAGAGGGCTTTCACG 965

Qy 821 GAAAGAACTGTTCCCAAGAATGCCAGTGCATAATGGAGGGACGTGTGATGCTGCCACAG 880
 | | | | | | | | | | | | | | | | | | | | | |
 Db 966 GACCCAAGTGTCTCCAGGAATGTCGCTGCCACAACGGCGGCCTCTGTGACCGATTCACTG 1025

Qy 881 GCCAATGTCATTGCAGTCCAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTTG 940
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1026 GGCAGTGCCGCTGCGCTCCGGGTACACTGGGGATCGGTGCCGGGAGGAGTGCCCGGTGG 1085

Qy 941 GGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTAC 1000
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1086 GCCGCTTTGGGCAGGACTGTGCTGAGACGTGCGACTGCGCCCCGGACGCCCGTTGCTTCC 1145

Qy 1001 ACGTGAGCGGCGCATGCCTCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCC 1060
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1146 CGGCCAACGGCGCATGTCTGTGCGAACACGGCTTCACTGGGGACCGCTGCACGGATCGCC 1205

Qy 1061 TGTGTCCTGAGGGGCTCTACGGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGA 1120
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1206 TCTGCCCCGACGGCTTCTACGGTCTCAGCTGCCAGGCCCCCGCACCTGCGACCGGGAGC 1265

Qy 1121 AACTCATAGCTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGAC 1180
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1266 ACAGCCTCAGCTGCCACCCGATGAACGGGGAGTGTCTCCTGCCTGCCGGGCTGGGCGGGCC 1325

Qy 1181 TCTACTGTAATGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCA 1240
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1326 TCCACTGCAACGAGAGCTGCCCCGAGGACACGCATGGGCCAGGGTGCCAGGAGCACTGTC 1385

Qy 1241 GCTGCCAAAATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGACCTGTGCCCCAGGAT 1300
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1386 TCTGCCTGCACGGTGGCGTCTGCCAGGCTACCAGCGCCTCTGTGAGTGCAGCGCGGGTT 1445

Qy 1301 TCAAAGGAATTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCT 1360
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1446 ACACGGGCCCTCACTGTGCTAGTCTTTGTCTCCTGACACCTACGGTGTCAACTGTTCTG 1505

Qy 1361 CTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCA 1420

Db	1506	CACGCTGCTCATGTGAAAAATGCCATCGCCTGCTCACCCATCGACGGCGAGTGCCTCTGCA	1565
Qy	1421	AGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGATGTCCCAGTGGCACATGGGGCTTTG	1480
Db	1566	AGGAAGGTTGGCAGCGTGGTAACTGCTCTGTGCCCTGCCCACCCGGAACCTGGGGCTTCA	1625
Qy	1481	GCTGTAACCTTAACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCT	1540
Db	1626	GTTGCAATGCCAGCTGCCAGTGTGCCCATGAGGCAGTCTGCAGCCCCCAAACCTGGAGCCT	1685
Qy	1541	GCACGTGTGCACCTGGATGGCGCGGGGAGAAATGCCAACTTCCCTGCCAGGATGGCACGT	1600
Db	1686	GTACCTGCACCCCTGGGTGGCATGGGGCCCACTGCCAGCTGCCCTGTCCGAAGGGGCAGT	1745
Qy	1601	ACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCA	1660
Db	1746	TTGGAGAAGGTTGTGCCAGTCGCTGTGACTGTGACCACTCTGATGGCTGTGACCCTGTTC	1805
Qy	1661	CGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTG	1720
Db	1806	ATGGACGCTGTCACTGCCAGGCTGGCTGGATGGGTGCCCGCTGCCACCTGTCTCTGCCCTG	1865
Qy	1721	AGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCT	1780
Db	1866	AGGGCTTATGGGGAGTCAACTGTAGCAACACCTGCACCTGCAAGAATGGGGGCACCTGTC	1925
Qy	1781	CCCTTGATGATGGCATCTGCGAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGAGAGGA	1840
Db	1926	TCCCTGAGAATGGCAACTGCGTGTGTGCGCCCGGATTCCGGGGCCCCCTCCTGCCAGAGAT	1985
Qy	1841	TCTGCTCCCCTGGTTTTTATGGGCATCGCTGCAGCCAGACATGCCCACAGTGCCTTCACA	1900
Db	1986	CCTGTCAGCCTGGCCGCTATGGCAAACGCTGTGTGC-----CCTGCAAGTGCCTAACC	2039
Qy	1901	GCAGCGGGCCCTGCCACCACATCACGGCCCTGTGTGACTGCTTGCCTGGCTTCACAGGCG	1960
Db	2040	ACTCC---TTCTGCCACCCCTCGAACGGGGCCTGCTACTGCCTGGCTGGCTGGACAGGCC	2096
Qy	1961	CCCTCTGCAATGAAGTGTGTCCAGTGGCAGATTTGGGAAAAACTGTGCAGGAATTTGTA	2020
Db	2097	CCGACTGCTCCCAGCCATGCCCTCCAGGACACTGGGGAGAAAACTGTGCCAGACCTGCC	2156
Qy	2021	CCTGCACCAACAACGGAACCTGTAACCCCATTGACAGATCTTGTGAGTGTACCCCGGTT	2080
Db	2157	AATGTCACCATGGTGGGACCTGCCATCCCAGGATGGGAGCTGTATCTGCCCCCTAGGCT	2216
Qy	2081	GGATTGGCAGTGAAGTGTCTCAACCATGTCCACCTGCCCCTGGGGCCCCAACTGCATCC	2140
Db	2217	GGACTGGACACCACTGCTTAGAAGGCTGCCCTCTGGGGACATTTGGTGCTAACTGCTCCC	2276
Qy	2141	ACACGTGCAACTGCCATAATGGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCA	2200
Db	2277	AGCCATGCCAGTGTGGTCTGGAGAAAAGTGCCACCCAGAGACTGGGGCCTGTGTATGTC	2336
Qy	2201	CTCCTGG 2207	

Db 2337 CCCCAGG 2343

RESULT 11

US-10-052-648A-5

; Sequence 5, Application US/10052648A
; Publication No. US20040005558A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Fuad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shimkets, Richard
; APPLICANT: Smithson, Glennda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER: US/10/052,648A
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2919

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-052-648A-5

Query Match 16.0%; Score 546.2; DB 15; Length 2919;
Best Local Similarity 55.3%; Pred. No. 8e-166;
Matches 1231; Conservative 0; Mismatches 903; Indels 93; Gaps 5;

```
Qy      62 GGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACT 121
      ||  ||  | ||| ||  ||  || || ||  || ||  || || || || ||
Db      129 GGCTGGCTGGAACCTCTCAACCCAGTGATCCCAATACCTGCAGCTTCTGGGAAAGCTTCA 188

Qy      122 CAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC- 180
      |  ||  | ||| ||  ||  || || ||  |  |  ||  | |||
Db      189 CTACCACCACCAAGGAGTCCCACTCCCGCCCTTCAGCCTGCTCCCTCAGAGCCCTGCG 248

Qy      181 --ACTGACATTCTAAACTGGTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCT 238
      |  ||  | ||| ||  ||  || || ||  ||  || || || ||  |
Db      249 AGCGGCCCTGGGAGGGCCCCCATACTTGCCCCAGCCACGGTTGTATACGGACCGTGT 308

Qy      239 ATCGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTGTCTGGATTTTATG 298
      | || ||  | || || || ||  |  |  || || || || || || || ||
Db      309 ACCGTGAGGTGGTGAAGACGGACCACCGCCAGCGCCTGCAGTGCTGCCATGGCTTCTATG 368

Qy      299 AAAGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAATGTGTCCATGGTCGCTGTATTG 358
      | ||| |||  | || || || || || || || || || || || || || ||
Db      369 AGAGCAGGGGGTTCTGTGTCCCGCTCTGTGCCAGGAGTGTGTCCATGGCCGTTGTGTGG 428

Qy      359 CTCCAAACACCTGTCAGTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCG 418
      | || ||  || || || || || || || || || || || || || || ||
Db      429 CACCCAATCAGTGCCAATGTGTGCCAGGCTGGCGGGGCGACGACTGTTCCAGTGAGTGTG 488

Qy      419 ATGGTGATCACTGGGGTCCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGT 478
      |  || || || || || || ||  | ||| |||  ||  ||
Db      489 CCCCAGGAATGTGGGGGCCACAGTGTGACAAGCCCTGCAGCTGCGGCAACAACAGCTCGT 548

Qy      479 GCAACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGG 538
      |  || |||  | || ||  ||  || || || || || || || || || ||
Db      549 GTGATCCCAAGAGTGGGGTATGTTCTTGCCCTTCTGGTCTGCAGCCCCCGAACTGCCTTC 608

Qy      539 ACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATG 598
      | || |||  | || || || || || || || || || || || || || ||
Db      609 AGCCCTGTACCCCTGGCTACTATGGCCCTGCCTGCCAGTTCGGCTGCCAGTGCC---ATG 665

Qy      599 GAGCCACCTGCGACCACGTACCGGGGAATGCCGCTGCCACACAGGATACACCGGAGCCT 658
      | || || || || || || || || || || || || || || || || || ||
Db      666 GGGCACCTGCGATCCCCAGACTGGAGCCTGCTTCTGCCCCGAGAGAGAACTGGGCCCA 725

Qy      659 TCTGTGAGGATCTTTGTCTCTCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTT 718
      || || || || || || || || || || || || || || || || || ||
Db      726 GCTGTGACGTGTCTGTTCAGGGCACTTCTGGCTTCTTCTGCCCCAGCACCCATTCTT 785

Qy      719 GTCAAAATGGAGGAGTGTGTATCAGCTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGA 778
      | || || || || || || || || || || || || || || || || || ||
Db      786 GCCAAAATGGAGGTGTCTTCAAACCCACAGGGCTCCTGCAGCTGCCCCCTGGCTGGA 845

Qy      779 T----- 779
```

Db	846	TGGTATGGAGGGTGGGGCCTGTGGGCATGGGGTGTGGGTCTGGGGAGAATTCTGTGGGTG	905
Qy	780	-----GGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTG	820
Db	906	GTGCTAAGCAGGGCTCCAAGGGCACCATCTGCTCCCTGCCCTGCCAGAGGGCTTTCACG	965
Qy	821	GAAAGAACTGTTCCCAAGAATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAG	880
Db	966	GACCCAACTGCTCCCAGGAATGTGCTGCCACAACGGCGGCCTCTGTGACCGATTCACTG	1025
Qy	881	GCCAATGTCATTGCAGTCCAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTG	940
Db	1026	GGCAGTGCCGCTGCGCTCCGGGTACACTGGGGATCGGTGCCGGGAGGAGTGCCCGGTG	1085
Qy	941	GGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTTACC	1000
Db	1086	GCCGCTTTGGGCAGGACTGTGCTGAGACGTGCGACTGCGCCCCGGACGCCCGTTGCTTCC	1145
Qy	1001	ACGTGAGCGGCGCATGCCTCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCC	1060
Db	1146	CGGCCAACGGCGCATGTCTGTGCGAACACGGCTTCACTGGGGACCGCTGCACGGATCGCC	1205
Qy	1061	TGTGTCCTGAGGGGCTCTACGGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGGAAA	1120
Db	1206	TCTGCCCCGACGGCTTCTACGGTCTCAGCTGCCAGGCCCGCCGACCTGCGACCGGGAGC	1265
Qy	1121	ACACTCATAGCTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGAC	1180
Db	1266	ACAGCCTCAGCTGCCACCCGATGAACGGGGAGTGCTCCTGCCCTGCCGGGCTGGGCGGGCC	1325
Qy	1181	TCTACTGTAATGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCA	1240
Db	1326	TCCACTGCAACGAGAGCTGCCCCGAGGACACGCATGGGCCAGGGTGCCAGGAGCGCTGTC	1385
Qy	1241	GCTGCCAAAATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGACCTGTGCCCCAGGAT	1300
Db	1386	TCTGCCTGCACGGTGGCGTCTGCCAGGCTACCAGCGGCCTCTGTCAGTGCGCGCCGGGTT	1445
Qy	1301	TCAAAGGAATTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCCCT	1360
Db	1446	ACACGGGCCCTCACTGTGCTAGTCTTTGTCTCCTGACACCTACGGTGTCAACTGTTCTG	1505
Qy	1361	CTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCA	1420
Db	1506	CACGCTGCTCATGTGAAAATGCCATCGCCTGCTCACCCATCGACGGCGAGTGCGTCTGCA	1565
Qy	1421	AGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTG	1480
Db	1566	AGGAAGGTTGGCAGCGTGGTAACTGCTCTGTGCCCTGCCACCCGGAACCTGGGGCTTCA	1625
Qy	1481	GCTGTAACCTTAACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCTGGACGGGACCT	1540
Db	1626	GTTGCAATGCCAGCTGCCAGTGTGCCCATGAGGCAGTCTGCAGCCCCAACTGGAGCCT	1685
Qy	1541	GCACGTGTGCACCTGGATGGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGT	1600


```

; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1976
; LENGTH: 1970
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (1890)..(1890)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (1970)..(1970)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1976

```

```

Query Match          15.8%; Score 540.6; DB 14; Length 1970;
Best Local Similarity 69.4%; Pred. No. 4.1e-164;
Matches 735; Conservative 0; Mismatches 324; Indels 0; Gaps 0;

```

```

Qy      917 GGTGCCAGGATGAGTGTCTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGT 976
        ||||| || ||||| || | ||| || | || | | || || || || ||
Db      1498 GGTGCCAAGAGGAGTGCCCTTCGGGTCCTTCGGCTTCCAGTGCTCACAGCGCTGTGACT 1439

Qy      977 GTGTCAACGGAGGGAAGTGTTACCACGTGAGCGGCGCATGCCCTCTGTGAAGCAGGCTTTG 1036
        | ||| || ||| ||||| | | | || ||| |||| | ||||
Db      1438 GCCACAATGGGGGGCAGTGTTACCCACCACGGGTGCCTGCGAGTGTGAGCCTGGCTACA 1379

Qy      1037 CTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGAGGGGCTCTACGGCATCAAATGTGACA 1096
        ||| ||||| | | || ||||| || ||||| || | ||| || |
Db      1378 AGGGCCCACGCTGCCAGGAGCGACTGTGCCCGGAGGGCCTGCATGGCCCAGGCTGCACCC 1319

Qy      1097 AACGGTGTCCCTGCCACTTGGAACAACTCATAGCTGTCACCCCATGTCTGGAGAGTGTG 1156
        | || ||||| || || ||||| |||| |||| | ||||| |||
Db      1318 TGCCCTGCCCCTGTGACGCTGACAACACCATCAGCTGCCACCCAGTAAGTGGAGCTTGTA 1259

Qy      1157 CCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAATGAGACATGTTCTCCTGGATTCTACG 1216
        ||||| |||| ||||| || | | |||| |||| | || || || | |||
Db      1258 CCTGCCAGCCAGGCTGGTCTGGTCACCACTGCAATGAATCCTGCCCTGTTGGCTACTATG 1199

Qy      1217 GGGAAAGCTTGCCAGCAGATCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACTG 1276
        | || | ||||| | |||| ||| || ||||| || ||||| |||| | ||||
Db      1198 GCGATGGCTGCCAGCTGCCTTGACCTGTGAGAATGGCGCCGACTGCCACAGCATCACTG 1139

Qy      1277 GAAAGTGACCTGTGCCCCAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTGG 1336
        | ||||| ||||| || || |||| ||| | || | || | || |
Db      1138 GGGGCTGCACTTGTGCTCCGGGCTTCATGGGAGAGGTCTGTGCCGTTTCCTGTGCAGCAG 1079

```


; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 113
; LENGTH: 5036
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (230)...(3379)
US-09-796-753-113

Query Match 15.4%; Score 528.6; DB 10; Length 5036;
Best Local Similarity 58.3%; Pred. No. 6.1e-160;
Matches 987; Conservative 0; Mismatches 694; Indels 12; Gaps 3;

```
Qy      515 GCTTCCGGGGCTGGCGCTGCGAGGACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTC 574
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      771 GTTCCAGTGCCCCGAAGTGCCTTCAGCCCTGTACCCCTGGCTACTATGGCCCTGCCTGCC 830

Qy      575 ATCAGAGATGCCAGTGCCAGAATGGAGCCACCTGCGACCACGTACGCGGGGAATGCCGCT 634
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      831 AGTTCGCTGCCAGTGCC---ATGGGGCACCTGCGATCCCCAGACTGGAGCCTGCTTCT 887

Qy      635 GCCCACCAGGATACACCGGAGCCTTCTGTGAGGATCTTTGTCCTCCTGGTAAACATGGTC 694
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      888 GCCCCGAGAGAGAACTGGGCCAGCTGTGACGTGTCTGTTCCAGGGCACTTCTGGCT 947

Qy      695 CACAGTGTGAGCAGAGATGCCCTTGTCAAATGGAGGAGTGTGTCATCACGTCACTGGAG 754
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      948 TCTTCTGCCCCAGCACCCATCCTTGCCAAATGGAGGTGTCTTCCAAACCCACAGGGCT 1007

Qy      755 AATGCTCTTGCCCTTCTGGCTGGATGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTC 814
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1008 CCTGCAGCTGCCCCCTGGCTGGATGGGCACCATCTGCTCCCTGCCCTGCCAGAGGGCT 1067

Qy      815 GCTTTGGAAAGAACTGTTCCCAAGAATGCCAGTGCCATAATGGAGGGACGTGTGATGCTG 874
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1068 TTCACGGACCCAAGTCTCCAGGAATGTCGCTGCCACAACGGCGGCCTCTGTGACCGAT 1127

Qy      875 CCACAGGCCAATGTCATTGCAGTCCAGGATACACAGGGGAACGGTGCCAGGATGAGTGTC 934
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1128 TCACTGGGCAGTGCCGCTGCGCTCCGGGTTACACTGGGGATCGGTGCCGGGAGGAGTGCC 1187

Qy      935 CTGTTGGGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGT 994
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1188 CGTGGGCGCTTTGGGGCAGGACTGTGCTGAGACGTGCGACTGCGCCCCGGACGCCCGTT 1247

Qy      995 GTTACCACGTGAGCGGCGCATGCCTCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAG 1054
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1248 GCTTCCCGGCCAACGGGCGCATGTCTGTGCGAACACGGCTTCACTGGGGACCGCTGCACGG 1307

Qy     1055 CACGCCTGTGTCTGAGGGGCTCTACGGCATCAAATGTGACAAACGGTGTCCCTGCCACT 1114
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1308 ATCGCTCTGCCCCGACGGCTTCTACGGTCTCAGCTGCCAGGCCCCCTGCACCTGCGACC 1367

Qy     1115 TGGAAAACACTCATAGCTGTACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGT 1174
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1368 GGGAGCACAGCCTCAGCTGCCACCCGATGAACGGGGAGTGCTCCTGCCTGCCGGGCTGGG 1427

Qy     1175 CAGGACTCTACTGTAATGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGA 1234
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1428 CGGGCTTCCACTGCAACGAGAGCTGCCCCGAGGACACGCATGGGCCAGGGTGCCAGGAGC 1487

Qy     1235 TCTGCAGCTGCCAAAATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGACCTGTGCCC 1294
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Db 1488 ACTGTCTCTGCCTGCACGGTGGCGTCTGCCAGGCTACCAGCGGCCTCTGTCACTGCGCGC 1547
 Qy 1295 CAGGATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACT 1354
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1548 CGGGTTACACGGGCCCTCACTGTGCTAGTCTTTGTCTCCTGACACCTACGGTGTCAACT 1607
 Qy 1355 GTTCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTA 1414
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1608 GTTCTGCACGCTGCTCATGTGAAAATGCCATCGCCTGCTCACCCATCGACGGCGAGTGCG 1667
 Qy 1415 CTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGATGTCCCAGTGGCACATGGG 1474
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1668 TCTGCAAGGAAGGTTGGCAGCGTGGTAACGTCTGTGCCCTGCCACCCGGAACCTGGG 1727
 Qy 1475 GCTTTGGCTGTAACTTAACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACG 1534
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1728 GCTTCAGTTGCAATGCCAGCTGCCAGTGTGCCCATGAGGCAGTCTGCAGCCCCCAAAGT 1787
 Qy 1535 GGACCTGCACGTGTGCACCTGGATGGCGCGGGGAGAAATGCCAACTTCCCTGCCAGGATG 1594
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1788 GAGCCTGTACCTGCACCCCTGGGTGGCATGGGGCCCACTGCCAGCTGCCCTGTCCGAAGG 1847
 Qy 1595 GCACGTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACC 1654
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1848 GGCAGTTTGGAGAAGGTTGTGCCAGTCGCTGTGACTGTGACCACTCTGATGGCTGTGACC 1907
 Qy 1655 CTACCACGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTGT 1714
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1908 CTGTTTCATGGACGCTGTCACTGCCAGGCTGGCTGGATGGGTGCCCGCTGCCACCTGTCCT 1967
 Qy 1715 GTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTT 1774
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1968 GCCCTGAGGGCTTATGGGGAGTCAACTGTAGCAACACCTGCACCTGCAAGAATGGGGGCA 2027
 Qy 1775 CATGCTCCCCTGATGATGGCATCTGCGAGTGTGCACCAGGCTTCCGAGGCACCACTTGTC 1834
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2028 CCTGTCTCCCTGAGAATGGCAACTGCGTGTGTGCACCCGGATTCCGGGGCCCCCTCCTGCC 2087
 Qy 1835 AGAGGATCTGCTCCCCTGGTTTTTATGGGCATCGCTGCAGCCAGACATGCCCACAGTGCG 1894
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2088 AGAGATCCTGTACAGCCTGGCCGCTATGGCAAACGCTGTGTGCC-----CTGCAAGTGCG 2141
 Qy 1895 TTCACAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGTGACTGCTTGCCTGGCTTCA 1954
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2142 CTAACCACTCCTT---CTGCCACCCCTCGAACGGGACCTGCTACTGCCTGGCTGGCTGGA 2198
 Qy 1955 CAGGCGCCCTCTGCAATGAAGTGTGTCCAGTGGCAGATTTGGGAAAAAAGTGTGCAGGAA 2014
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2199 CAGGCCCCGACTGCTCCCAGCCATGCCCTCCAGGACACTGGGGAGAAAAAGTGTGCCCAGA 2258
 Qy 2015 TTTGTACCTGCACCAACAACGGAACCTGTAACCCATTGACAGATCTTGTCACTGTTACC 2074
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2259 CCTGCCAATGTCAACCATGGTGGGACCTGCCATCCCAGGATGGGAGCTGTATCTGCCCCC 2318
 Qy 2075 CCGGTTGGATTGGCAGTGACTGCTCTCAACCATGTCCACCTGCCCCTGGGGCCCCAACT 2134
 | | | | | | | | | | | | | | | | | | | | | |
 Db 2319 TAGGCTGGACTGGACACCACTGCTTAGAAGGCTGCCCTCTGGGGACATTTGGTGCTAACT 2378

Qy 2135 GCATCCACACGTGCAACTGCCATAATGGAGCTTTCTGCAGCGCCTACGATGGGGAATGTA 2194
 || ||| | ||| | || | |||| ||| | | | |||| |||
 Db 2379 GCTCCCAGCCATGCCAGTGTGGTCCTGGAGAAAAGTGCCACCCAGAGACTGGGGCCTGTG 2438
 Qy 2195 AATGCACTCCTGG 2207
 ||| | || ||
 Db 2439 TATGTCCCCCAGG 2451

RESULT 14

US-09-833-381-1910/c
 ; Sequence 1910, Application US/09833381
 ; Patent No. US20020132090A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Robison, Keith E.
 ; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
 ; FILE REFERENCE: 5800-119
 ; CURRENT APPLICATION NUMBER: US/09/833,381
 ; CURRENT FILING DATE: 2001-04-11
 ; PRIOR APPLICATION NUMBER: 09/516,448
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 2050
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 1910
 ; LENGTH: 5197
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(5197)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-833-381-1910

Query Match 15.3%; Score 524.6; DB 9; Length 5197;
 Best Local Similarity 58.0%; Pred. No. 1.3e-158;
 Matches 972; Conservative 0; Mismatches 694; Indels 11; Gaps 2;

Qy 531 CTGCGAGGACCGCTGTGAGCAGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTG 590
 |||| | |||| | ||| ||||| | ||| || | ||||| |||
 Db 4243 CTGCCTTCAGCCCTGTACCCCTGGCTACTATGGCCCTGCCTGCCAGTTCGCTGCCAGTG 4184
 Qy 591 CCAGAATGGAGCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCACCAGGATACAC 650
 || |||| || ||||| || | || || | ||| ||||| ||| ||
 Db 4183 CC---ATGGGGCACCCCTGCGATCCCAGACTGGAGCCTGCTTCTGCCCCGAGAGAGAAC 4127
 Qy 651 CGGAGCCTTCTGTGAGGATCTTTGTCCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAG 710
 || || ||||| | ||| | | || | ||| || || |
 Db 4126 TGGGCCCAGCTGTGACGTGTCCTGTTCCCAGGGCACTTCTGGCTTCTTCTGCCCCAGCAC 4067
 Qy 711 ATGCCCTTGTCAAAATGGAGGAGTGTGTATCACGTCACTGGAGAATGCTCTTGCCCTTC 770
 |||| | ||||| || | || | || || ||| |||| |
 Db 4066 CCATCCTTGCCAAAATGGAGGTGTCTTCAAACCCACAGGGCTCCTGCAGCTGCCCCC 4007
 Qy 771 TGGCTGGATGGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTGGAAAGAACTG 830
 ||||| ||||| || || | ||| |||| |||| ||| |||||
 Db 4006 TGGCTGGATGGGCACCATCTGCTCCCTGCCCTGCCAGAGGGCTTTCACGGACCCAACTG 3947

Qy 831 TTCCAAGAATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAGGCCAATGTCA 890
 ||||| ||||| | ||||| || || || ||||| ||| || || || |
 Db 3946 CTCCCAGGAATGTCGCTGCCACAACGGCGGCCTCTGTGACCGATTCACTGGGCAGTGCCG 3887

Qy 891 TTGCAGTCCAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTTGGGACCTATGG 950
 ||| ||| || ||||| ||||| ||||| ||| ||||| || || || || || ||
 Db 3886 CTGCGCTCCGGGTACACTGGGGATCGGTGCCGGGAGGAGTGCCCGGTGGGCCGCTTTGG 3827

Qy 951 CGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACCACGTGAGCGG 1010
 ||||| ||||| ||| | || | | | | || | || | | |||
 Db 3826 GCAGGACTGTGCTGAGACGTGCGACTGCGCCCCGACGCCCGTGTCTTCCCGGCCAACGG 3767

Qy 1011 CGCATGCCTCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCGAAGCACGCCTGTGTCTGA 1070
 ||||| || || || || ||||| ||||| || ||||| || ||||| || || || || ||
 Db 3766 CGCATGTCTGTGCGAACACGGCTTCACTGGGGACCGCTGCACGGATCGCCTCTGCCCCGA 3707

Qy 1071 GGGGCTCTACGGCATCAAATGTGACAAACGGTGTCCTTGCCACTTGAAAACACTCATAG 1130
 || ||||| || || || | | || ||||| || ||| ||| | ||
 Db 3706 CGGCTTCTACGGTCTCAGCTGCCAGGCCCCCTGCACCTGCGACCGGGAGCACAGCCTCAG 3647

Qy 1131 CTGTCACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGACTCTACTGTAA 1190
 ||| ||||| || || ||||| ||||| ||||| ||| || ||| ||||| ||
 Db 3646 CTGCCACCCGATGAACGGGGAGTGCTCCTGCCTGCCGGGTGGGCGGGCCTCCACTGCAA 3587

Qy 1191 TGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAAAA 1250
 |||| || || | | | || || || ||||| || ||| ||||| |
 Db 3586 CGAGAGCTGCCCCGAGGACACGCATGGGCCAGGTGCCAGGAGCACTGTCTCTGCCTGCA 3527

Qy 1251 TGGGGCAGACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGATTCAAAGGAAT 1310
 || | || || | | | || || || || || || || || || ||
 Db 3526 CGGTGGCGTCTGCCAGGCTACCAGCGGCCTCTGTCACTGCGCGCCGGGTACACGGGGCC 3467

Qy 1311 TGA CTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCTCGCTGTGG 1370
 | |||| ||| | || |||| | ||||| || | ||||| || |||||
 Db 3466 TCACTGTGCTAGTCTTTGTCTCTGACACCTACGGTGTCAACTGTTCTGCACGCTGCTC 3407

Qy 1371 CTGTAAAAATGATGCAGTCTGCTCTCCTGTGGACGGGTCTTGTACTTGCAAGGCAGGCTG 1430
 ||| ||||| | ||||| || | ||||| || ||||| ||| ||
 Db 3406 ATGTGAAAATGCCATCGCCTGCTCACCCATCGACGGCGAGTGCGTCTGCAAGGAAGGTTG 3347

Qy 1431 GCACGGGGTGGACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTGGCTGTAACTT 1490
 ||| | | ||||| | || || || || || ||||| || || ||
 Db 3346 GCACCGTGGTAACCTGCTCTGTGCCCTGCCACCCGGAACCTGGGGCTTCAGTTGCAATGC 3287

Qy 1491 AACATGCCAGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGTGC 1550
 | ||||| || | | || || ||||| || || || || || || ||
 Db 3286 CAGCTGCCAGTGTGCCCATGAGGCAGTCTGCAGCCCCAACTGGAGCCTGTACCTGCAC 3227

Qy 1551 ACCTGGATGGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCACGTACGGGCTGAA 1610
 ||||| |||| |||| | || | || |||| | | || || || ||
 Db 3226 CCCTGGGTGGCATGGGGCCCACTGCCAGCTGCCCTGTCCGAAGGGGAGTTTGGAGAAGG 3167

Qy 1611 CTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCACGGGCCATTG 1670
 ||||| |||| |||| |||| | ||||| |||| || || || ||
 Db 3166 TTGTGCCAGTCGCTGTGACTGTGACCACTCTGATGGCTGTGACCCTGTTTCATGGACGCTG 3107

Qy 1671 CCGCTGCCTCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTGAGGGACGCTG 1730
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 3106 TCAGTGCCAGGCTGGCTGGATGGGTGCCCCTGCCACCTGTCTGCCCTGAGGGCTTATG 3047

Qy 1731 GGGCCCCAACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGA 1790
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 3046 GGGAGTCAACTGTAGCAACACCTGCACCTGCAAGAATGGGGGCACCTGTCTCCCTGAGAA 2987

Qy 1791 TGGCATCTGCGAGTGTGCACCAGGCTTCCGAGGCACCACTTGTGAGAGGATCTGCTCCCC 1850
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2986 TGGCAACTGCGTGTGTGCACCCGGATTCCGGGGCCCCCTCCTGCCAGAGATCCTGTGAGCC 2927

Qy 1851 TGGTTTTTATGGGCATCGCTGCAGCCAGACATGCCACAGTGCCTTACAGCAGCGGGCC 1910
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2926 TGGCCGCTATGGCAAACGCTGTG-----TGCCCTGCAAGTGCCTAACCCTCCTTC 2875

Qy 1911 CTGCCACCACATCACCGGCTGTGTGACTGCTTGCCTGGCTTACAGGCGCCCTCTGCAA 1970
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2874 TGCCACCCCTTCGAACGGGACCTGCTACTGCCTGGCTGGCTGGACAGGCCCCGACTGCTC 2815

Qy 1971 TGAAGTGTGTCCCAGTGGCAGATTTGGGAAAAACTGTGCAGGAATTTGTACCTGCACCAA 2030
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2814 CCAGCCATGCCCTCCAGGACACTGGGGAGAAAACTGTGCCCAGACCTGCCAATGTACCA 2755

Qy 2031 CAACGGAACCTGTAACCCCATTGACAGATCTTGTCAGTGTTACCCCGGTTGGATTGGCAG 2090
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2754 TGGTGGGACCTGCCATCCCAGGATGGGAGCTGTATCTGCCCCCTAGGCTGGACTGGACA 2695

Qy 2091 TGAAGTGTGTCCCAGTGGCAGATTTGGGAAAAACTGTGCAGGAATTTGTACCTGCACCAA 2150
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2694 CCACTGCTTAGAAGGCTGCCCTCTGGGGACATTTGGTGCTAACTGCTCCCAGCCATGCCA 2635

Qy 2151 CTGCCATAATGGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATGCACTCCTGG 2207
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 2634 GTGTGGTCCTGGAGAAAAGTGCCACCCAGAGACTGGGGCCTGTGTATGTCCCCCAGG 2578

RESULT 15

US-10-052-648A-1

; Sequence 1, Application US/10052648A

; Publication No. US20040005558A1

; GENERAL INFORMATION:

; APPLICANT: Anderson, David
 ; APPLICANT: Burgess, Catherine
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Colman, Steven
 ; APPLICANT: Edinger, Shlomit R.
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: MacDougall, John R.
 ; APPLICANT: Mehraban, Fuad
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Rothenberg, Mark
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Smithson, Glennda

```
; APPLICANT:  Spytek, Kimberly A.
; APPLICANT:  Stone, David J.
; APPLICANT:  Vernet, Corine A.M.
; APPLICANT:  Zerhusen, Bryan D.
; TITLE OF INVENTION:  PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION:  USING THE SAME
; FILE REFERENCE:  21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER:  US/10/052,648A
; CURRENT FILING DATE:  2002-12-09
; PRIOR APPLICATION NUMBER:  60/262,454
; PRIOR FILING DATE:  2001-01-18
; PRIOR APPLICATION NUMBER:  60/272,920
; PRIOR FILING DATE:  2001-03-02
; PRIOR APPLICATION NUMBER:  60/284,549
; PRIOR FILING DATE:  2001-04-18
; PRIOR APPLICATION NUMBER:  60/303,229
; PRIOR FILING DATE:  2001-07-05
; PRIOR APPLICATION NUMBER:  60/262,892
; PRIOR FILING DATE:  2001-01-19
; PRIOR APPLICATION NUMBER:  60/263,605
; PRIOR FILING DATE:  2001-01-23
; PRIOR APPLICATION NUMBER:  60/269,098
; PRIOR FILING DATE:  2001-02-15
; PRIOR APPLICATION NUMBER:  60/264,159
; PRIOR FILING DATE:  2001-01-25
; PRIOR APPLICATION NUMBER:  60/265,517
; PRIOR FILING DATE:  2001-01-31
; PRIOR APPLICATION NUMBER:  60/271,855
; PRIOR FILING DATE:  2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS:  97
; SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH:  3063
;   TYPE:  DNA
;   ORGANISM:  Homo sapiens
US-10-052-648A-1
```

```
Query Match          14.9%;  Score 511;  DB 15;  Length 3063;
Best Local Similarity  55.2%;  Pred. No. 2.4e-154;
Matches 1164;  Conservative    0;  Mismatches  850;  Indels    93;  Gaps    5;
```

```
Qy      62 GGACAGCATCACCTCTGAATCTTGAAGACCCTAATGTGTGTAGCCACTGGGAAAGCTACT 121
      ||  ||  | |||| || |  || || ||  || ||| ||||| |||| |
Db      47 GGCTGGCTGGAACCTCTCAACCCAGTGATCCCAATACCTGCAGCTTCTGGGAAAGCTTCA 106

Qy     122 CAGTGACTGTGCAAGAGTCATACCCACATCCCTTTGATCAAATTTACTACACGAGCTGC- 180
      |  ||  | |||| || | | ||||  |  |  ||  | ||||
Db     107 CTACCACCACCAAGGAGTCCCACTCCCGCCCCTTCAGCCTGCTCCCTCAGAGCCCTGCG 166

Qy     181 --ACTGACATTCTAAACTGGTTTTAAATGCACGCGGCACAGAGTCAGCTATCGGACAGCCT 238
      |  |  |  || ||| | | || | ||  || |||| | |
Db     167 AGCGGCCCTGGGAGGGCCCCCATACTTGCCCCAGCCACGGTTGTATACGGACCGTGT 226

Qy     239 ATCGACATGGGGAGAAGACTATGTATAGGCGCAAGTCTCAGTGTTCCTGGATTTTATG 298
      | || || | || ||||  | | |  |||| || | ||| || ||||
Db     227 ACCGTCAGGTGGTGAAGACGGACCACCGCCAGCGCCTGCAGTGCTGCCATGGCTTCTATG 286
```

Qy	299	AAAGCGGGGAAATGTGTGTCCCCCACTGTGCTGATAAAATGTGTCCATGGTCGCTGTATTG	358
Db	287	AGAGCAGGGGGTTCTGTGTCCCGCTCTGTGCCAGGAGTGTGTCCATGGCCGTTGTGTGG	346
Qy	359	CTCCAAACACCTGTCACTGTGAGCCTGGCTGGGGAGGGACCAACTGCTCCAGTGCCTGCG	418
Db	347	CACCCAATCAGTGCCAATGTGTGCCAGGCTGGCGGGGCGACGACTGTTCCAGTGAGTGTG	406
Qy	419	ATGGTGATCACTGGGGTCCCCACTGCACCAGCCGGTGCCAGTGCAAAAATGGGGCTCTGT	478
Db	407	CCCCAGGAATGTGGGGGCCACAGTGTGACAAGCCCTGCAGCTGCGGCAACAACAGCTCGT	466
Qy	479	GCAACCCCATCACCGGGGCTTGCCACTGTGCTGCGGGCTTCCGGGGCTGGCGCTGCGAGG	538
Db	467	GTGATCCCAAGAGTGGGGTATGTTCTTGCCCTTCTGGTCTGCAGCCCCGAAGTGCCTTC	526
Qy	539	ACCGCTGTGAGCAGGGGCACCTATGGTAACGACTGTCATCAGAGATGCCAGTGCCAGAATG	598
Db	527	AGCCCTGTACCCCTGGCTACTATGGCCCTGCCTGCCAGTTCGCTGCCAGTGCC---ATG	583
Qy	599	GAGCCACCTGCGACCACGTACGGGGGAATGCCGCTGCCCACCAGGATACACCGGAGCCT	658
Db	584	GGGCACCTGCGATCCCCAGACTGGAGCCTGCTTCTGCCCCGAGAGAGAACTGGGCCCCA	643
Qy	659	TCTGTGAGGATCTTTGTCTCTCCTGGTAAACATGGTCCACAGTGTGAGCAGAGATGCCCTT	718
Db	644	GCTGTGACGTGTCTCTGTTCCAGGGCACTTCTGGCTTCTTCTGCCCCAGCACCCATTCTT	703
Qy	719	GTCAAAATGGAGGAGTGTGTTCATCACGTCACTGGAGAATGCTCTTGCCCTTCTGGCTGGA	778
Db	704	GCCAAATGGAGGTGTCTTCCAAACCCACAGGGCTCCTGCAGCTGCCCCCTGGCTGGA	763
Qy	779	T-----	779
Db	764	TGGTATGGAGGGTGGGGCCTGTGGGCATGGGGTGTGGGTCTGGGGAGAATTCTGTGGGTG	823
Qy	780	-----GGGCACAGTGTGTGGTCAGCCTTGCCCCGAGGGTCGCTTTG	820
Db	824	GTGCTAAGCAGGGCTCCAAGGGCACCATCTGCTCCCTGCCCTGCCAGAGGGCTTTCACG	883
Qy	821	GAAAGAACTGTTCCCAAGAATGCCAGTGCCATAATGGAGGGACGTGTGATGCTGCCACAG	880
Db	884	GACCCAATGCTCCCAGGAATGTCGCTGCCACAACGGCGGCCTCTGTGACCGATTCACTG	943
Qy	881	GCCAATGTCATTGCAGTCCAGGATACACAGGGGAACGGTGCCAGGATGAGTGTCTGTG	940
Db	944	GGCAGTGCCGCTGCGCTCCGGGTTACACTGGGGATCGGTGCCGGGAGGAGTGCCCGGTG	1003
Qy	941	GGACCTATGGCGTTCTCTGTGCTGAGACCTGCCAGTGTGTCAACGGAGGGAAGTGTACC	1000
Db	1004	GCCGCTTTGGGCAGGACTGTGCTGAGACGTGCGACTGCGCCCCGGACGCCCGTTGCTTCC	1063
Qy	1001	ACGTGAGCGGCGCATGCCTCTGTGAAGCAGGCTTTGCTGGCGAGCGCTGCCAAGCACGCC	1060
Db	1064	CGGCCAACGGCGCATGTCTGTGCGAACACGGCTTCACTGGGGACCGCTGCACGGATCGCC	1123

Qy 1061 TGTGTCCTGAGGGGCTCTACGGCATCAAATGTGACAAACGGTGTCCCTGCCACTTGAAAA 1120
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1124 TCTGCCCCGACGGCTTCTACGGTCTCAGCTGCCAGGCCCTGCACCTGCGACCGGGAGC 1183

Qy 1121 ACACCTCATAGCTGTACCCCATGTCTGGAGAGTGTGCCTGCAAGCCGGGCTGGTCAGGAC 1180
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1184 ACAGCCTCAGCTGCCACCCGATGAACGGGGAGTGTCTCTGCCTGCCGGGCTGGGCGGGCC 1243

Qy 1181 TCTACTGTAATGAGACATGTTCTCCTGGATTCTACGGGGAAGCTTGCCAGCAGATCTGCA 1240
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1244 TCCACTGCAACGAGAGCTGCCCCGAGGACACGCATGGGCCAGGGTGCCAGGAGCACTGTC 1303

Qy 1241 GCTGCCAAAATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGGAT 1300
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1304 TCTGCCTGCACGGTGGCGTCTGCCAGGCTACCAGCGCCTCTGTCTAGTGCAGCGCCGGGTT 1363

Qy 1301 TCAAAGGAATTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGATAAACTGTTCTCT 1360
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1364 ACACGGGGCCCTCACTGTGCTAGTCTTTGTCTCTGACACCTACGGTGTCAACTGTTCTG 1423

Qy 1361 CTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCTCTGTGGACGGGTCTTGTACTTGCA 1420
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1424 CACGCTGCTCATGTGAAATGCCATCGCCTGCTACCCATCGACGGCGAGTGCCTCTGCA 1483

Qy 1421 AGGCAGGCTGGCACGGGGTGGACTGCTCCATCAGATGTCCAGTGGCACATGGGGCTTTG 1480
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1484 AGGAAGGTTGGCAGCGTGGTAAGTCTGTGCCCTGCCACCCGGAACCTGGGGCTTCA 1543

Qy 1481 GCTGTAACCTAACATGCCAGTGCCTCAACGGGGAGCCTGCAACACCCCTGGACGGGACCT 1540
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1544 GTTGCAATGCCAGCTGCCAGTGTGCCATGAGGCAGTCTGCAGCCCCCAAACCTGGAGCCT 1603

Qy 1541 GCACGTGTGCACCTGGATGGCGCGGGGAGAAATGCCAACTTCCCTGCCAGGATGGCACGT 1600
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1604 GTACCTGCACCCCTGGGTGGCATGGGGCCCACTGCCAGTGCCTGTCCGAAGGGGAGT 1663

Qy 1601 ACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCTGCCACCCTACCA 1660
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1664 TTGGAGAAGGTTGTGCCAGTGCCTGTGACTGTGACCACTCTGATGGCTGTGACCCTGTTC 1723

Qy 1661 CGGGCCATTGCCGCTGCCTCCCGGGATGGTCAGGTGTCCACTGTGACAGCGTGTGTGCTG 1720
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1724 ATGGACGCTGTCTAGTGCCAGGCTGGCTGGATGGGTGCCCGCTGCCACCTGTCTCTGCCCTG 1783

Qy 1721 AGGGACGCTGGGGCCCCAACTGCTCCCTGCCCTGCTACTGTAAAAATGGGGCTTCATGCT 1780
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1784 AGGGCTTATGGGGAGTCAACTGTAGCAACACCTGCACCTGCAAGAAATGGGGGACCTGTC 1843

Qy 1781 CCCCTGATGATGGCATCTGCGAGTGTGCACAGGCTTCCGAGGCACCACTTGTCTAGAGGA 1840
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1844 TCCCTGAGAATGGCAACTGCGTGTGTGCACCCGATTCGGGGCCCCCTCTGCCAGAGAT 1903

Qy 1841 TCTGCTCCCCTGGTTTTTATGGGCATCGCTGCAGCCAGACATGCCCACAGTGCCTTCACA 1900
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1904 CCTGTCAGCCTGGCCGCTATGGCAAACGCTGTGTGC-----CCTGCAAGTGCAGTAACC 1957

Qy 1901 GCAGCGGGCCCTGCCACCACATCACCGGCCTGTGTGACTGCTTGCTGGCTTCACAGGCG 1960

```

      | |      ||||| | | | | | | | | | | | |
Db    1958 ACTCC---TTCTGCCACCCCTCGAACGGGACCTGCTACTGCCTGGCTGGCTGGACAGGCC 2014

Qy    1961 CCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTTGGGAAAACTGTGCAGGAATTTGTA 2020
      || ||| | | | | | | | | | | | | | |
Db    2015 CCGACTGCTCCCAGCGCTGCCCTCTGGGGACATTTGGTGCTAACTGCTCCCAGCCATGCC 2074

Qy    2021 CCTGCACCAACAACGGAACCTGTAACCCCATTTGACAGATCTTGTCAGTGTTACCCCGGTT 2080
      || | | | | | | | | | | | | | |
Db    2075 AGTGTGGTCCTGGAGAAAAGTGCCACCCAGAGACTGGGGCCTGTGTATGTCCCCCAGGGC 2134

Qy    2081 GGATTGG 2087
      | | |
Db    2135 ACAGTGG 2141

```

Search completed: March 30, 2004, 08:56:09
Job time : 1167 secs